



University of Zurich  
Zurich Open Repository and Archive

Winterthurerstr. 190  
CH-8057 Zurich  
<http://www.zora.uzh.ch>

---

*Year: 2007*

---

## Automated generalisation in digital and mobile cartography

Burghardt, D

Burghardt, D. Automated generalisation in digital and mobile cartography. 2007, University of Zurich, Faculty of Science.

Postprint available at:  
<http://www.zora.uzh.ch>

Posted at the Zurich Open Repository and Archive, University of Zurich.  
<http://www.zora.uzh.ch>

Originally published at:  
University of Zurich, Faculty of Science, 2007.

# Automated generalisation in digital and mobile cartography

## Abstract

The variety of cartographic presentations of spatial and non spatial information has been significantly extended through web mapping and applications for mobile devices. The habilitation thesis in hand presents current research results on the automated derivation of digital, cartographic presentations with a major focuss on automated generalisation. The introductory sections describe challenges of modern cartography embedded in the context of geovisualisation and geoinformation sciences. Furthermore, current developments are shown in relation to the conventional topographic map production. These are characterised through a transition from paper map production towards GIS based cartography utilising multirepresentation databases. With that, prerequisites are given for an efficient data update process and the flexible consideration of different usage scenarios and user requirements. Exemplary applications of location-based services are presented and special requirements on mobile cartography are explained. The main part of this habilitation thesis is dedicated to research on the area of automated generalisation. During the definition of terms, generalisation is explained by the concept of abstraction. With that the task of automated generalisation is considered in a wider sense, besides the preservation of legibility constraints the major focuss is set on purpose orientation and user specific derivation of cartographic presentation. The discussion of previous generalisation research is carried out by a comparison of selected conceptual and technical frameworks. Furthermore, a categorisation and analysis is made of research presented at Generalisation Workshops of the ICA Commission on Generalisation and Multiple Representation to illustrate current and future research topics. The research publications integrated in this habilitation thesis cover several thematic topics. They present new methods and approaches which support both automated production of topographic maps as well as cartographic presentation on mobile devices. At the beginning it will be shown that the concept of Webservices can be applied successfully to offer generalisation functionality independent of platform in a modular and hierarchically structured form. Following that methods are proposed for the modelling and analysis of cartographic requirements through constraints. Besides the formalised description of object properties the explicit modelling of relations between object representations are also essential for a successful automated generalisation. In this thesis two new generalisation algorithms will be introduced on the basis of different optimisation methods (snakes and mesh simplification). In particular the typification approach supports the generalisation of topographic maps, e.g. the typification of buildings, but can also be applied within mobile applications such as for the conflict free presentation of Points of Interest. Finally research question are investigated in the context of mobile information systems. A process chain is used to geo-enable spatial and non-spatial data making them available for locationbased services. Real-time requirements for cartographic presentation within this context can be satisfied through the usage of hierarchical data structures. In the subsequent discussion results and limitations of the presented approaches are worked out in a holistic view. There it is shown, that selected solutions can be transferred from the traditional cartographic domain to mobile cartography. The outlook enumerates possible future research topics, induced through the extended application of new technologies such as mobile information systems, positioning techniques, service oriented architectures, interaction possibilities and multimedia. New developments attempt on extending the standardised, universal information exchange on the basis of topographic maps with specialised, user centered presentation of spatial and non-spatial information.

# **Automated generalisation in digital and mobile cartography**

**Habilitation**

vorgelegt der  
Mathematisch-naturwissenschaftlichen Fakultät  
der  
Universität Zürich

von

**Dirk Burghardt**

aus Deutschland

Zürich, 2007





## Abstract

The variety of cartographic presentations of spatial and non spatial information has been significantly extended through web mapping and applications for mobile devices. The habilitation thesis in hand presents current research results on the automated derivation of digital, cartographic presentations with a major focuss on automated generalisation.

The introductory sections describe challenges of modern cartography embedded in the context of geovisualisation and geoinformation sciences. Furthermore, current developments are shown in relation to the conventional topographic map production. These are characterised through a transition from paper map production towards GIS based cartography utilising multirepresentation databases. With that, prerequisites are given for an efficient data update process and the flexible consideration of different usage scenarios and user requirements. Exemplary applications of location-based services are presented and special requirements on mobile cartography are explained.

The main part of this habilitation thesis is dedicated to research on the area of automated generalisation. During the definition of terms, *generalisation* is explained by the concept of abstraction. With that the task of automated generalisation is considered in a wider sense, besides the preservation of legibility constraints the major focuss is set on purpose orientation and user specific derivation of cartographic presentation. The discussion of previous generalisation research is carried out by a comparison of selected conceptual and technical frameworks. Furthermore, a categorisation and analysis is made of research presented at *Generalisation Workshops* of the ICA Commission on Generalisation and Multiple Representation to illustrate current and future research topics.

The research publications integrated in this habilitation thesis cover several thematic topics. They present new methods and approaches which support both automated production of topographic maps as well as cartographic presentation on mobile devices. At

the beginning it will be shown that the concept of Webservices can be applied successfully to offer generalisation functionality independent of platform in a modular and hierarchically structured form. Following that methods are proposed for the modelling and analysis of cartographic requirements through *constraints*. Besides the formalised description of object properties the explicit modelling of relations between object representations are also essential for a successful automated generalisation.

In this thesis two new generalisation algorithms will be introduced on the basis of different optimisation methods (*snakes* and *mesh simplification*). In particular the typification approach supports the generalisation of topographic maps, e.g. the typification of buildings, but can also be applied within mobile applications such as for the conflict free presentation of Points of Interest. Finally research question are investigated in the context of mobile information systems. A process chain is used to *geo-enable* spatial and non-spatial data making them available for location-based services. Real-time requirements for cartographic presentation within this context can be satisfied through the usage of hierarchical data structures.

In the subsequent discussion results and limitations of the presented approaches are worked out in a holistic view. There it is shown, that selected solutions can be transferred from the traditional cartographic domain to mobile cartography. The outlook enumerates possible future research topics, induced through the extended application of new technologies such as mobile information systems, positioning techniques, service oriented architectures, interaction possibilities and multimedia. New developments attempt on extending the standardised, universal information exchange on the basis of topographic maps with specialised, user centered presentation of spatial and non-spatial information.

## Zusammenfassung

Das Spektrum kartographischer Präsentation räumlicher und nicht räumlicher Informationen hat sich durch Anwendungen im Web und auf mobilen Endgeräten stark verbreitert. Die vorliegende Arbeit präsentiert aktuelle Forschungsergebnisse der automatischen Ableitung digitaler, kartographischer Präsentationen mit einem Schwerpunkt auf der automatischen Generalisierung.

In den einführenden Abschnitten wird zunächst auf die Herausforderungen der modernen Kartographie im Kontext von Geovisualisierung und Geoinformationswissenschaften eingegangen. Ausserdem werden die aktuellen Entwicklungen im Bereich der amtlichen Kartenproduktion aufgezeigt. Diese sind gekennzeichnet durch einen Wechsel von reiner Papierkartenproduktion hin zu GIS-basierter Kartographie unter Nutzung von Multirepräsentationsdatenbanken. Damit existiert einerseits die Voraussetzung für eine effektive Datenfortführung und andererseits kann flexibel auf unterschiedliche Anwendungsszenarien und Nutzerbedürfnisse reagiert werden. Beispielhaft werden hierfür Anwendung orts-basierter Dienste (Location-based Services) vorgestellt und spezielle Anforderungen der mobilen Kartographie erläutert.

Der Hauptteil der Arbeit beschäftigt sich mit der Forschung auf dem Gebiet der automatischen Generalisierung. In der Begriffsdefinition zur *Generalisierung* wird die Abstraktion als wesentliche Komponente hervorgehoben. Die Aufgaben der automatischen Generalisierung sind damit bewusst weiter gefasst, neben der Einhaltung von Lesbarkeitsbedingungen liegt der Schwerpunkt auf einer zweckorientierten und nutzerspezifischen Ableitung kartographischer Präsentationen. Die Aufarbeitung der bisherigen Generalisierungsforschung erfolgt durch einen Vergleich ausgewählter Generalisierungskonzepte und -plattformen. Ausserdem werden die Themenschwerpunkte der Internationalen Workshops der ICA Kommission für Generalisierung und Multirepräsentation analysiert, um gegenwärtige und zukünftige Forschungsschwerpunkte zu illustrieren.

Die in dieser Arbeit zusammengefassten Forschungspublikationen lassen sich unterschiedlichen Themenschwerpunkten zuordnen. Dabei werden Methoden und Ansätze vorgestellt, welche sowohl die automatische Produktion topographischer Karten als auch die kartographische Präsentation auf mobilen Geräten unterstützen. Zunächst wird gezeigt, wie das Konzept der Webservices erfolgreich eingesetzt werden kann, um Generalisierungsfunktionalität plattformunabhängig, modular und hierarchisch gegliedert zur Verfügung zu stellen. Anschliessend werden Ansätze zur Modellierung und Analyse kartographischer Bedingungen (*constraints*) vorgestellt. Neben der formalisierten Beschreibung von Objekteigenschaften ist die explizite Modellierung von Beziehung zwischen Objektrepräsentationen essentiell für eine erfolgreiche automatische Generalisierung.

In der Arbeit werden zwei neue Generalisierungsalgorithmen auf der Basis unterschiedlicher Optimierungsmethoden (*snakes* und *mesh simplification*) vorgestellt. Speziell der Typifikationsansatz ermöglicht Anwendungen sowohl in der Generalisierung topographischer Karten, z.B. für Gebäude, als auch in mobilen Anwendungen, z.B. für die konfliktfreie Darstellung von Points of Interest. Abschliessend werden Forschungsfragen im Kontext mobiler Informationssysteme vorgestellt. Es wird gezeigt, wie räumliche und nicht räumliche Daten für ortsbasierte Dienste zur Verfügung gestellt werden können (*geo-enabling*). Echtzeitanforderungen in der kartographischen Darstellung sind hierbei durch den Einsatz hierarchischer Datenstrukturen realisierbar.

In der nachfolgenden Diskussion werden die Ergebnisse und Grenzen der vorgestellten Ansätze in einer gesamtheitlichen Betrachtung herausgearbeitet. Dabei zeigt sich, dass ausgewählte Lösungsansätze aus dem Bereich der traditionellen Kartographie auch auf Fragestellungen der mobilen Kartographie übertragbar sind. Im Ausblick werden mögliche zukünftige Forschungsschwerpunkte aufgezählt, induziert durch die verbreitete Anwendung neuer Technologien wie mobile Informationssysteme, Positionierungstechniken, Serviceorientierte Architekturen, Interaktionsmöglichkeiten und Multimedia. Neue Entwicklungen versuchen hierbei die standardisierte, universelle Informationsvermittlung auf der Basis topographischer Karten um spezialisierte, nutzerzentrierte Präsentation räumlicher und nicht räumlicher Information zu erweitern.

## Acknowledgments

I am most grateful to:

Robert Weibel for his support of my research over the last years, the fruitful discussions and the continuous input on this work. It is a great pleasure to work at his department.

Ross Purves and Alistair Edwardes for the endless corrections of my improvable English. Alistair especially for the cooperation within the WebPark project, which is reflected in this thesis through paper 6 and 7. Thanks Ross for your insights in teaching and research.

Matthias Bobzien und Ingo Petzold for the open atmosphere and the intensive cooperation within the DRIVE and SerAx projects. We managed to balance the disperse requirements of University and privat industry as shown in paper 3. Thanks also to Moritz Wittensöldner supporting the project with his programming and analytical abilities.

Moritz Neun und Stefan Steiniger for the successful research carried out within your PhD studies as part of DEGEN project. Thanks to Moritz for the "Services" presented in paper 1 and Stefan for the discussions on principal component analysis as shown in paper 2.

Alessandro Cecconi and Martin Galanda for the cooperation during the GENDEM project. Alessandro for the experimental work carried out with the mesh simplification approach presented in paper 5.

Ajay Mathur and the Team from Axes Systems AG for confidence in our ideas and the cooperation within the DRIVE and SerAx projects.

Finally I'd like to thank my wife Cornelia and the kids - Annabella, Amelie and Amanuel - for giving me a life beside the work.

The EU project WebPark was supported by the Swiss State Secretariat for Education and Research within the Fifth Framework Programme of the European Commission (Project No. IST-2000-31041)

The research within the DRIVE/SerAx project was supported by the Swiss Commission for Technology and Innovation (Projects No. 6559.3 ENS-ES and No. 7921.1 ESPP-ES)

The research within the DEGEN project was supported by the Swiss National Science Foundation (Projects No. 200020-101798 and No. 200020-116302)

The research within the GENDEM project was supported by the Swiss National Science Foundation (Project No. 20-59249.99)

# Contents

<b>List of Figures</b> .....	<b>xi</b>
<b>1 Introduction</b> .....	<b>1</b>
<b>2 Digital cartography</b> .....	<b>9</b>
2.1 Cartography, Geovisualisation and GIScience .....	9
2.2 Production cartography .....	13
2.3 Mobile cartography and location-based services .....	16
<b>3 Theory of automated map generalisation</b> .....	<b>23</b>
3.1 Abstraction and generalisation .....	23
3.2 Continuous generalisation .....	25
3.3 Conceptual frameworks for automated generalisation .....	34
3.4 Technical frameworks for automated generalisation .....	38
3.5 Generalisation state-of-the-art .....	42
3.6 Funded projects and research context .....	53
<b>4 Generalization Services on the Web</b> .....	<b>59</b>
4.1 Motivation .....	60
4.2 Web services for spatial applications and generalization ....	61
4.3 Application of service concepts to generalization .....	62
4.4 Characteristics of web-based generalization services .....	63
4.5 Generalization services categories .....	68
4.6 Implementation .....	72
4.7 Conclusion .....	78
<b>5 Usage of PCA in Automated Generalisation</b> .....	<b>81</b>
5.1 Introduction .....	82
5.2 Constraint space and Principal Component Analysis .....	82
5.3 Homogeneity of groups - subspace of constraint space ....	87
5.4 Detecting settlement types from building dataset .....	90
5.5 Detecting special buildings .....	97
5.6 Conclusion .....	99

<b>6</b>	<b>MRDB with Explicitly Modelled Relations</b>	<b>101</b>
6.1	Introduction	102
6.2	Theory of horizontal, vertical and update relations	103
6.3	Similarities and Differences	113
6.4	Prototype and Results	116
6.5	Conclusion	119
<b>7</b>	<b>Controlled Line Smoothing by Snakes</b>	<b>121</b>
7.1	Introduction	122
7.2	Energy minimization for smoothing of line objects	124
7.3	Automated control of parameter selection	126
7.4	Typification	132
7.5	Conclusion	135
<b>8</b>	<b>Mesh Simplification for Building Typification</b>	<b>137</b>
8.1	Introduction	138
8.2	Theory of mesh optimization	138
8.3	Mesh simplification adapted for building typification	140
8.4	Discussion of results and possible improvements	152
8.5	Conclusions	154
<b>9</b>	<b>Geo-enabling Data for Mobile Information</b>	<b>155</b>
9.1	Introduction	156
9.2	Context of research	156
9.3	Evaluation of environmental and geographic information	157
9.4	Overview of data transformation and management	157
9.5	Remodelling	160
9.6	Geo-enabling	163
9.7	Design considerations for visualisation	165
9.8	Conclusions	169
<b>10</b>	<b>Automated Generalisation of Dynamically Generated Maps</b>	<b>171</b>
10.1	Introduction	172
10.2	Constraints for generalisation of symbolised point objects	173
10.3	Approaches for conflict free presentation of point objects	174
10.4	WebPark project - mobile tourist guide for recreational areas	180
<b>11</b>	<b>Conclusion</b>	<b>183</b>
11.1	Discussion of the papers presented	183
11.2	Outlook and further work	187



## List of Figures

1.1	Overview of chapters and related projects. . . . .	2
2.1	Perspective view from TomTom NAVIGATOR (left) and hybrid presentations from Microsoft Virtual Earth (right). . . . .	9
2.2	Classification of web maps (Kraak and Brown, 2001). . . . .	10
2.3	Major requirements of the users for applications, content and provision of geospatial information (Fornfeld et al., 2003). . . . .	13
2.4	Planned derivation process of Digital Cartographic Models from Topographic Landscape Model at swisstopo (Feldmann and Kreiter, 2006). . . . .	14
2.5	Iterative process of map creation, derived from Kriz (2001). . . . .	16
2.6	Structuring on non-spatial information in the Webpark mobile information system. User interface with geographical filter. . . . .	19
2.7	Examples of different symbols for different target groups (Patalaviciute et al., 2005). . . . .	21
3.1	Seville 2004, 9 panels each 32" x 32", oils on used plywood (left); Avignon 2004, 48" x 32", oils and oil pastel on used plywood (right) by Alice Dodge. . . . .	23
3.2	Portrayal schemes showing differing degrees of spatial relationships (Edwardes et al., 2005c). . . . .	25
3.3	Discontinuity introduced by abstraction and generalisation. . . . .	26
3.4	The principle of a morphing transformation is to create an intermediate state out of two 'keyframes' represented by LOD <sub>25</sub> and LOD <sub>200</sub> (Cecconi, 2003). . . . .	28
3.5	The principle of progressive transmission of vector data (Yang et al., 2007). . . . .	29
3.6	Fish eye projection with linear radial projection (urban - left, rural - right). The examples on the bottom are generalised through the selection of streets and buildings (Rappo, 2003). . . . .	32
3.7	Transformation by Laplacian smoothing (Edwardes, 2007). . . . .	33

3.8	Generalisation process of dynamically generated maps (Edwardes et al., 2003b). . . . .	34
3.9	The Portrayal model and generalisation model adapted from Cuthbert (1998) and following the generalisation typology of Weibel and Dutton (1998). . . . .	35
3.10	A conceptual framework of automated map generalisation (Brassel and Weibel, 1988). . . . .	36
3.11	Possible sequencing of generalisation tasks (Hardy and Lee, 2005). . . . .	37
3.12	General architecture of an agent (Duchêne, 2004). . . . .	38
3.13	Generalisation Service of GiMoDig, internal workflow. The numbers show the order in which the process steps are carried out. The branches a and b indicate alternative routes through the process (Lehto, 2003). . . . .	39
3.14	Client/server architecture for real-time generalisation based on XSLT extended by Java classes (Mannes, 2004). . . . .	39
3.15	Service framework of the commercial cartographic GIS expand (Axes Systems AG) with generalisation and registry server and the workflow management system (Petzold et al., 2006b). . . . .	42
4.1	Generalization services acting as a middleware (above) or as a research platform. . . . .	63
4.2	A registry for generalization services. . . . .	73
4.3	Three ways to implement process communication. . . . .	74
4.4	Available services (from the Generalization Service registry). . . . .	76
4.5	Algorithm interface generated dynamically from an interface description. . . . .	78
5.1	Evaluation of states with constraint severity (left) and equivalent presentation with standardised (middle) and reduced constraint space (right). . . . .	84
5.2	Left - "Principal Axis Transformation" with simple rotation to create a better adapted coordinate system; middle - reduction of two correlated components ( <i>area</i> , <i>perimeter</i> ) into one principal component ( <i>size</i> ); right - correlation between <i>area</i> and <i>shape</i> index ( $\frac{perimeter}{2\sqrt{\pi \cdot area}}$ ), due to the fact that smaller buildings are generally simpler than larger ones. . . . .	86
5.3	Mean, minimal and maximal deviation values of 16 measures for 40 alignments. . . . .	88
5.4	Visualised homogeneity of building alignments. Inhomogeneous alignments have lighter grey values. . . . .	89

5.5	Part of the training data (left), where every group of buildings represent one settlement type. Classification results for the same data are shown in the picture on the right side. Correctly classified buildings have the same grey value as in the left picture. Data: VECTOR25, reproduced by permission of Swisstopo (BA057008). . . . .	93
5.6	Transformed and projected training data from 7 dimensional property space (defined by the measures) onto an artificial property plane. The settlement type regions were defined manually. . . . .	94
5.7	Obtaining more homogeneous settlement type regions by application of a spatial median filter. Left picture: original classification; right picture: after application of the filter. Data: VECTOR25, reproduced by permission of Swisstopo (BA057008). . . . .	96
5.8	Result of the identification of special buildings, for separate treatment during map generalisation. Data: VECTOR25, reproduced by permission of Swisstopo (BA057008). . . . .	98
6.1	Example of generalized map features in different resolutions (scales 1 : 25 000, 1 : 50 000 and 1 : 100 000) and corresponding symbolization. Figures taken from SSC (2005), with permission from the Swiss Society of Cartography. . . . .	102
6.2	Horizontal relations on the example of building alignments: The horizontal relation has two elements, one connecting the buildings <i>a</i> and <i>b</i> , and the other connecting the buildings <i>c</i> , <i>d</i> and <i>e</i> . . . . .	105
6.3	Examples of horizontal relations - buffer area around a generalization zone (left) and line buffering (right). . . . .	106
6.4	The detail selection of a river network (original, left) without consideration of structural relations (middle) and with explicit focus on dominant orientation (right). Figures taken from SSC (2005), with permission from Swiss Society of Cartography. . . . .	108
6.5	Schematic view on the vertical relation. . . . .	109
6.6	Schematic view of an update relation. . . . .	110
6.7	Combination of horizontal relation (blue), vertical relation (red) and update relation (green). . . . .	111
6.8	Combination of horizontal relations and vertical relation. Left: The vertical relation elements are connected by a horizontal relation. Right: The building alignments (horizontal relations, blue) in different resolutions are connected by a vertical relation (red). . . . .	112

6.9	Two different modelings of a vertical relation: modeling by an $n$ -ary relation (left) and by a binary relation (right). . . . .	114
6.10	Two different modelings of an update relation: modeling by an $n$ -ary relation (left) and by a binary relation (right). . . . .	115
6.11	UML-Diagram of the implementation of relations, reduced to the main classes and attributes. An open arrow represents a directed association, with given multiplicity. A closed arrow represents a specialization. . . . .	116
6.12	Horizontal Relations: (a) original situation (b) partitioning by trans-hydro-graph (c) building alignments (d) neighborhood relation between buildings and nearest road. . . . .	117
6.13	Vertical relation between features of different resolutions with scale dependent symbolization (left) and update relation between features in one resolution (right). . . . .	118
6.14	Interference of horizontal, vertical and update relation. Situations: (a) original, (b) updated, (c) generalized original, (d) updated and generalized. Compare Figure 6.12 and Figure 6.13. The order of the screenshots is analogue to the order in Figure 6.7. . . . .	118
7.1	Levels of detail for a line object at different scales, after Töpfer (1974) . . . . .	123
7.2	Excessive shifting of end points of line objects depending on the degree of smoothing. . . . .	126
7.3	Forced fixed boundary points as a result of introducing multiple end points. . . . .	127
7.4	Forced fixed boundary points as a result of using a number of duplicated points at the line end rotated by $180^\circ$ around the crossroads. Left - Duplicated and rotated points at the line end. Right - Resulting lines after smoothing. . . . .	127
7.5	Calculation of number of duplicated points depending on distances between original and smoothed line to achieve fixed boundaries. . . . .	128
7.6	Segmentation based on intersection points between original and smoothed line. Left - Arrows show the intersection points. Right - Intersection points as a measure of sinuosity. . . . .	129
7.7	Concatenation of segments with different sinuosity attribute. . . . .	130
7.8	An example of the use of snakes for line smoothing. Smoothing was applied after line segmentation (left - original lines, right - smoothed lines). . . . .	130
7.9	Calculation of vertices (white) and intersection points (gray). . . . .	132

7.10	Construction of new curve bends through translation and exaggeration. ....	133
7.11	Typification examples of roads (to compare the manual generalized topographic maps are shown in the background - © 2004 swisstopo (BA046257)). ....	134
7.12	Typification and smoothing of road network. ....	134
7.13	Frequency dependent typification with different amounts of preprocessed smoothing. ....	135
8.1	Example of reducing the number of vertices of a mesh (Hoppe et al., 1993). ....	139
8.2	Flowchart for the typification operator of the generalization process for the feature class building, carried out by the mesh simplification technique. The highlighted part on the left computes the position of the new placeholders, while the right part defines the representation. ....	141
8.3	The method of edge collapsing for mesh simplification (left). On the right a sequence of edge collapses is shown. $v_i$ describes the center of gravity of building $b_i$ . ....	141
8.4	Phases of the mesh simplification technique for the step 'position'. ....	142
8.5	The positions of the placeholders for the following scales: a) buildings in LOD <sub>25</sub> , b) 1:25'000 (original data set), c) 1:50'000, d) 1:75'000, e) 1:100'000 and f) 1:200'000 with the objects' center of gravity in LOD <sub>200</sub> (all not to scale). Data: VECTOR25/200 © swisstopo (BA034957). ....	144
8.6	Interpolation function based on an extension of Töpfer's radical law with two parameters - number of objects ( $nb_{200}$ ) at a given second scale (left) and the power $1/p$ to modify the increase of the curve (right). ....	145
8.7	Mesh simplification based on Delaunay triangulation removing shortest edges between centroids (left) or edges with shortest distance between the objects outline (right). ....	146
8.8	Weighting of Delaunay edges to influence distribution of objects. ....	146
8.9	The positions of the placeholders with $s_u = 0.0$ (left) and $s_u = 1.0$ (right), both not to scale. Data: VECTOR25 © swisstopo (BA034957). ....	148
8.10	Comparison of unweighted object selection (left) with weighted selection considering absolute size (middle) and size differences (right). ....	149

8.11	The positions and dimensions of the placeholders for different scales: a) 1:25'000 (representing LOD25), b) 1:50'000, c) 1:75'000, d) 1:100'000, e) 1:200'000 and f) 1:200'000 (representing LOD200) (all at scale 1:25'000). Data: VECTOR25/200 © swisstopo (BA034957). . . . .	151
8.12	Typification of buildings computed for different scales with correct proportions. For scale 1:200'000 the computed (a)) and the original (b)) representations are shown. Data: VECTOR25/200 © swisstopo (BA034957). . . . .	152
8.13	Left: The positions of the placeholders for scale 1:75'000 (not to scale). Right: The dimensions of the placeholder computed out for scale 1:75'000. . . . .	153
9.1	Data processing flow in WebPark. . . . .	159
9.2	Hierarchical stream basins and visibility regions. . . . .	163
9.3	Information model for WebPark. . . . .	165
9.4	Individual ungulate observations. . . . .	166
9.5	Lower part: Example point set and corresponding quadtree (note the count stored at each node). Upper part: Successive generalisations of deer observations. . . . .	167
9.6	Idealised Stahler stream ordering and tessellation using catchments at three levels showing progressive levels of zoom. . . . .	168
9.7	Comparison of hierarchical generalisation techniques using quadtree (left) and hierarchical stream ordering and associated catchments (right). . . . .	169
10.1	Ungeneralised thematic maps with places of interest (left) and animal observation data (right). . . . .	173
10.2	Reduction of symbol overlaps with clustering approaches (left: presentation without generalisation, right: result of clustering). . . . .	176
10.3	Usage of quadtree data structure for adaptive zooming. . . . .	177
10.4	Removal of symbol overlaps (left) with the quadtree approach (middle) and clustering methods (right). . . . .	178
10.5	Mosaic subdivision on the basis of the digital terrain model and hierarchical stream basins (catchments). . . . .	179
10.6	Comparison of hierarchical generalisation techniques with the original observation data. Left: Using quadtree. Right: Using hierarchical stream ordering and associated stream basins. . . . .	180
10.7	Group with mobile walking guide at the Swiss National Park. . . . .	181
10.8	Improvement of information access based on changes of the graphical user interface. . . . .	182

# 1 Introduction

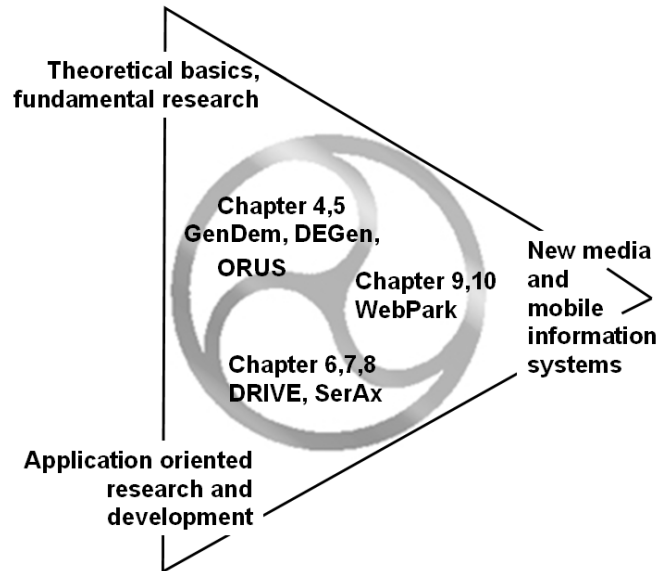
The aim of this habilitation thesis is to summarise five years of research in the domain of digital and mobile cartography. Recent developments in cartography have been characterised by rapid technological developments with enormous consequences for both geospatial information access, map use and of course, map production. The research focus of this work is centred on automated map generalisation, the search for a better understanding of the cartographic generalisation process and thus the possibility to improve automated solutions.

The ongoing and completed research projects described within this work lie at the intersection between fundamental research and application oriented developments. Furthermore, it is necessary to investigate both generic and specific solutions to consider the different requirements for paper map production versus cartographic presentation on new media and especially on mobile devices. New forms of mapping offer the possibilities to query not only individual layers of topographic data, but to ask questions of associated thematic data, which provides challenges in delivering cartographically appropriate representations.

The structure of the thesis reflects three main areas, theoretical basics and fundamental research (I), application oriented research and development (II) and new media and mobile information systems (III). The link connecting these areas lies in investigating the role of automated generalisation to these different areas. Figure 1.1 gives an overview of the research papers selected for this habilitation (Chapter 4-10) and illustrates their linkage within the working areas represented by the different research projects.

**Figure 1.1**

*Overview of chapters and related projects.*



The inner circle of Fig. 1.1 indicates the interrelationship of the research and development work carried out in the different projects:

- ❑ GenDem (May 2000- November 2002) - Map generalisation for thematic and on-demand mapping in GIS
- ❑ DEGen (May 2004 - September 2007) - Data enrichment for automated map generalisation
- ❑ ORUS (February 2006 - January 2009) - Ontology-driven recognition of urban structures from spatial databases
- ❑ WebPark (October 2001 - September 2004) - Location-based services in natural areas
- ❑ DRIVE (May 2004 - October 2005) - Derivation of digital vector models
- ❑ SerAx (November 2005 - October 2007) - Web based generalisation services for online maps

Theoretical approaches are transferred into practical solutions and, vice versa, problems from the domains of cartographic map production and geographical information access via mobile devices are explored. Subthemes of this work include the development of



on-demand and on-the-fly generalisation approaches, solutions for paper and mobile maps, the presentation of topographic data versus specific thematic information or the consideration of individual user requirements in contrast to precalculated generic solutions.

**Chapter 2** presents a characterisation of cartographic research in connection with geovisualisation and in the broader context of GIScience (Section 2.1). Furthermore, current tasks and challenges are described from the twin perspectives of production (Section 2.2) and mobile cartography (Section 2.3).

Section 2.1 summarises new mapping possibilities offered to cartography through user interaction, dynamic presentation and the application of multimedia information. In particular, the exploration of geospatial data through cartographic or map-related techniques has extended cartographic research in the direction of geovisualisation, especially when considered from a GIScience perspective. Basic research in the generalisation domain is closely related to exploration of geospatial data and the identification of important patterns, trends and relationships between features. An important requirement for automated generalisation is therefore the preservation of such information. Furthermore conventional data models must be enriched by explicitly modelling such patterns and relationships as are found within the data.

In Section 2.2 current challenges in production cartography are discussed as carried out at National Mapping Agencies (NMAs). These organisations are working on the introduction and extension of GIS-based mapping solutions with the aim of introducing flexible, seamless and user oriented mapping applications. In this way, NMAs can provide access to digital data as well as traditional paper maps containing selected layers and adapted symbolisation. Prerequisite include the utilisation of Multi-Representation Databases (MRDBs), which support the management of digital vector models with different measurement accuracies. Both automated and interactive generalisation methods are required for the derivation of digital landscape models and digital topographic maps.

Mobile cartography, a relatively new branch of cartography, is introduced in Section 2.3. The main difference between map presentations on mobile devices and traditional paper mapping stem from the possibility of considering device position derived from for example, GPS, from the small display size of the portable devices and from the consideration of context information for map adaptation. A very effective method of adapting content is the selec-

tion and presentation of information with respect to user position - so called "location-based services". In a preparation phase geo-enabling methods have to be applied to associate spatial and non-spatial information. The term "egocentric maps" is used to describe this new map type, characterised through short-term usage by individual users. Related research presented in this habilitation thesis concentrates on the development of real-time methods and the utilisation of hierarchical data structures to satisfy requirements for the dynamic, user oriented map portrayal.

In **Chapter 3** the term *generalisation* is introduced in the broader context of *abstraction*. Following that is a discussion about possible approaches to control abstraction levels in cartographic presentation. A distinction is made between an abstraction of features and attributes on one hand and abstraction of spatial relationships on the other.

In Section 3.2 requirements for continuous generalisation are presented to support infinitely variable zoom in web mapping and mobile applications. Generalisation operators such as simplification, elimination, collapse, aggregation or typification introduce discontinuities into a map sequence of varying scales. These abrupt changes can cause orientation difficulties. In research papers several methods have been proposed to realise smooth, interpolated transitions, for example, moving, morphing, fading, and appearing. To speed up the interpolation, a Multi-Representation Database can be used, which supports a server side management of precalculated representations at different resolutions. Furthermore continuous zooming can be combined with the progressive transmission of vector map data to reduce the waiting period on the client.

Continuous shape deformation methods are important for dynamic mapping and mobile applications. Application examples exist within the construction of area cartograms, fish-eye projection on small screens, shape similarity measures and functional grid approaches and their utilisation for cartographic generalisation. Detailed explanations of two possible methods - *active contour models* and *mesh simplification*, will be given here (Chapters 7 and 8, respectively). A main advantage of this deformation method is the continuous sequence of results which can be derived with one transformation.

Section 3.3 discusses important conceptual frameworks for automated generalisation. A coarse subdivision is made between model and cartographic generalisation, connected by a organisa-

tion/interpretation component in the case of dynamic mapping applications. The web mapping model by Cuthbert (1998) is presented and enriched with relevant generalisation components. The early framework introduced by Brassel and Weibel (1988) is the first comprehensive view of the generalisation process including structure recognition, process recognition, process modelling, process execution and final map display. Furthermore workflow and constraint based frameworks are presented to illustrate recent approaches to developing generalisation strategies. Section 3.4 contains a description of implementation frameworks that were developed in the context of research described in subsequent papers.

Section 3.5 presents a historical overview of the roots of generalisation research. Following that current research challenges are worked out based on an analysis of "hot topics" in recent research workshops and journal issues. The Section finishes with a listing of commercial available generalisation software and an enumeration of books and special issues related to automated generalisation research. Section 3.6 gives an overview of the research context through a short summary of funded projects and lead over to the collection of research papers.

The **research papers** in this habilitation thesis have been selected to illustrate the extended spectrum of automated generalisation research investigated based on the new possibilities and requirements identified for cartographic visualisation. These papers demonstrate generalisation concepts and solutions for different cartographic application areas, including topographic map production, web mapping and mobile applications. Furthermore the research presented gives insights into the multiple aspects of the generalisation process.

In the first paper about *Generalisation services* a modern, flexible architecture is proposed utilising the web service concept to delivering generalisation functionality both on-the-fly without any user interaction as well as in an interactive toolbox scenario with major control on the user side.

Burghardt, D., Neun, M. and Weibel, R. (2005). Generalization Services on the Web - a Classification and an Initial Prototype Implementation. *Cartography and Geographic Information Science*, Vol. 32, No. 4, pp. 257-268.

Chapter 4  
(first paper)

The second paper introduces a constraint space model for the generalised treatment of cartographic constraints. Based on this model *principal component analysis* can be utilised to determine the correlations and interdependencies of the constraints. This provides a basis for the selection of generalisation operators during the generalisation process.

Chapter 5  
(second paper)

Burghardt, D. and Steiniger, S. (2005). Usage of Principal Component Analysis in the Process of Automated Generalisation. *Proceedings of 22<sup>nd</sup> International Cartographic Conference, A Coruña, Spain*, CD-ROM.

Paper three investigates the enrichment of Multi-Representation Databases with explicitly modelled *horizontal*, *vertical* and *update relations*. The relations are further characterisations of features and their changes across scales and time. Thus, meta information relevant for initial and incremental generalisation can be stored.

Chapter 6  
(third paper)

Bobzien, M., Burghardt, D., Petzold, I., Neun, M. and Weibel, R. (2007). Multi-Representation Databases with Explicitly Modelled Horizontal, Vertical and Update Relations. *Cartography and Geographic Information Science*.

Papers four and five present new algorithms for the generalisation operations *smoothing* and *typification*. In particular, typification operators based on mesh simplification provide generalisation functionality for both topographic mapping and mobile applications.

Chapter 7  
(fourth paper)

Burghardt, D. (2005). Controlled Line Smoothing by Snakes. *GeoInformatica*, Vol. 9, No. 3, 2005, pp. 237-252.

Chapter 8  
(fifth paper)

Burghardt, D. and Alessandro Cecconi, A. (2007). Mesh Simplification for Building Typification. *International Journal of Geographical Information Science*, Volume 21, No. 3-4, 2007, pp. 283-298.

Papers six and seven are dedicated to the application of automated generalisation in the context of mobile devices. They investigate the special requirements which exist for generalisation and cartographic visualisation on small screens in real-time.

Paper six describes the *geo-enabling* and modelling process of spatial and non-spatial data to make them available for location-based services. The final paper explains two different approaches utilising *hierarchical data structures* for the real-time visualisation of cartographic icons.

Edwardes, A., Burghardt, D., Dias, E., Purves, R. and Weibel, R. (2005). Geo-enabling Spatially Relevant Data for Mobile Information Use and Visualisation. *Lecture Notes in Computer Science*. 5<sup>th</sup> International Workshop on Web and Wireless Geographical Information Systems, Springer-Verlag, pp. 78-92.

Chapter 9  
(sixth paper)

Burghardt, D., Edwardes, A., Purves, R. and Weibel, R. (2005). Automated generalisation of dynamic generated maps for mobile devices. *Kartographische Nachrichten* (originally in German), Vol. 55, No. 3, pp. 119-125.

Chapter 10  
(seventh paper)

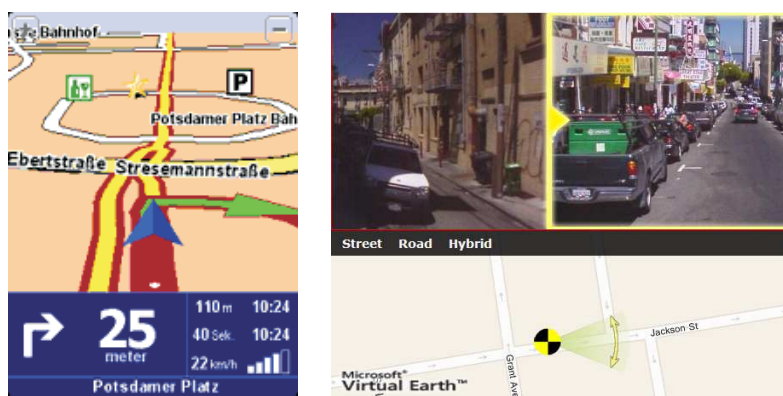
In the **Conclusion** chapter relevance and contributions of the papers selected for this habilitation are shown within the research concept introduced above. Also a categorisation of the papers is made in a broader research context. Therefore the categories of Table 3.1 and Table 3.2 from Chapter 3 are referred to, which cover recent and perspective research topics on automated generalisation and multiple representation. Open research questions are discussed within the application areas of *traditional topographic mapping* on the one hand and *user centered adaptive mapping* on the other. Thus the habilitation thesis concludes with challenges and possible trends for future research.



## 2 Digital cartography

### 2.1 Cartography, Geovisualisation and GIScience

**Challenges in digital cartography** In recent time, cartography and map presentations have evolved faster than ever. This evolution was initiated through the change from analogue to digital cartography, with enormous consequences on map use and map creation. Furthermore, with the availability of Geographic Information Systems (GIS), suddenly "cartography has gone from a read-only medium into being a read/write medium" (Dodson, 2005). The utilisation of the world wide web has offered new possibilities for map use and exchange of spatial data. Applications such as Google Earth / Google Maps or Windows Live Local offer mapping products that combine topographic information with aerial and satellite imagery. Car navigation systems such as TomTom NAVIGATOR present cartographic information successfully in spoken form and with perspective views (Fig. 2.1, left).



**Figure 2.1**  
Perspective view from TomTom NAVIGATOR (left) and hybrid presentations from Microsoft Virtual Earth (right).

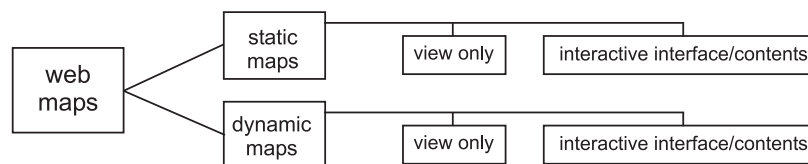
Figure 2.1 (right) shows an example from Microsoft Virtual Earth in which a 2D map presentation relates to the picture presentation of the surroundings. As the position in the map changes,

also the photographic presentation above is adapting. Intuitive graphical user interfaces and fast data access encourage people to use this kind of maps in their daily life. Kraak (2006, p. 128) points out that "in the past professional and public users lived in two separate worlds . . . today, the public and professional mapping worlds come even closer together."

In the digital age completely new mapping capabilities are offered to the users through the opportunity of *interaction* with the map (Andrienko and Andrienko, 1999; Dykes et al., 2005) and the application of *multimedia* information (Cartwright et al., 1999, 2004). Additionally animated *dynamic presentation* facilities are applicable especially interesting for the presentation of time dependent phenomena (DiBiase et al., 1992; Dorling, 1992; Kraak et al., 1997). Disadvantages of digital mapping products are the limitations by the size of the output device. Compensations are possible through interaction tools such as zooming, panning or the query for additional information.

A categorisation for web mapping applications distinguishing between static and dynamic maps on one side and the degree of interaction on the other side is shown in Figure 2.2 (Kraak and Brown, 2001). The supplementary information can be presented in various forms such as text, picture, video or sound. For example the incorporation of audio information into maps is now commonplace in car navigation systems. It presents guidance information in a spoken form. More information about the incorporation of multimedia information into a mobile information system and the effort to make it spatially explorable is given in Chapter 9 (Research Paper 6).

**Figure 2.2**  
Classification of web  
maps (Kraak and  
Brown, 2001).



The technological development has fundamentally changed the acquisition, management, analysis and representation of geospatial data. Based on that cartography has been extended with new research fields on geographic visualisation or short geovisualisation. Here the main focus is on exploration, the search for "unknowns", through cartographic or map-related techniques. The following definition of geovisualisation (GVis) is based on expla-



nations and comments from the web page of the ICA Commission on Visualization and Virtual Environments:

*Geovisualisation* is as a form of map-based information communication that emphasizes development and assessment of visual methods designed to facilitate the exploration, analysis, synthesis and presentation of georeferenced information.

*Geovisualisation*  
(GVis)

In geovisualisation there is a combined interest in both development of theory, tools, and methods of new map forms as well as in understanding how the tools and methods are used to prompt visual thinking. Kraak (2006, p. 129) reflects the intensions of GVis by saying that "the map is no longer the map as many of us know it" and stresses the function of a map "as flexible interface to geospatial data" that encourages exploration and identification of geospatial patterns and relationships. The role of maps in spatial knowledge acquisition is pointed out by MacEachren (1991, 1994b) as well. The transition from traditional maps to modern geovisualisation tools is characterised by new methods and applications to explore, synthesise, present and analyse data (Longley et al., 2005). Selected research issues in the field of geovisualisation are summarised by Kraak and MacEachren (2005):

- ❑ develop new forms of geographic information access and manipulation based on experiential and multi-sensory representation technologies
- ❑ develop extensible methods and tools that enable understanding of large and complex geospatial data sets
- ❑ develop geovisualisation methods and tools that support group work
- ❑ develop a human centred approach to geovisualisation

The exploratory part related to spatial phenomena has developed an own research direction of *exploratory data analysis* (Tukey, 1977), in recent time called *visual analytics* as the science of analytical reasoning facilitated by interactive visual interfaces. A current research and development agenda for visual analytics is published by Thomas and Cook (2005). Main topics are the development of analytical reasoning techniques, transformation and interaction techniques, as well as techniques to support presentation and dissemination of analysis results. The close relationship between cartographic user interactions and visual analytics can be seen in the

*Visual analytics*

*Visual information-seeking mantra* *Visual Information-Seeking Mantra* (Shneiderman, 1996), which summarises many visual design guidelines.

“Overview first, zoom and filter, then details on demand.”

There is the believe of this author that methods from scientific and information visualisation, developed for the presentation of interactive or animated digital images to users, are adaptable also for vector based cartographic portrayal.

**Digital cartography and GIScience** The extension of the cartographic spectrum becomes obvious through the name changes of cartographic associations and journals. For instance the “American Cartographer” was renamed in 1990 to “Cartography and Geographic Information Systems” and in 1999 to “Cartography and Geographic Information Science”. In Germany the “Kartographische Nachrichten” now carries the subtitle “Fachzeitschrift für Geoinformation und Visualisierung”. The International Cartographic Association has recently decided to add “The International Society for Cartography and Geographic Information Science” as a subtitle to its name (Kraak, 2005). Today cartographic research and development can be seen as an essential part of geovisualisation (Gahegan, 2005; Hearnshaw and Unwin, 1994), which itself is part of the broader domain of Geographical Information Science (short: GIScience). The term GIScience was introduced by Goodchild (1992) as “the science behind the systems”, dedicated to advancing our understanding of geographic processes and spatial relationships through improved theory, methods, technology, and data (Board on Earth Sciences and Resources (BESR), 2006). As McMaster and User (2003) say, a close relationship exists between GIS and numerous disciplines, including cartography, photogrammetry, geodesy, surveying, computer and information science, and statistics, among others. Scientists are in agreement on the term “geographic information science (GIScience)” to describe the theory behind these fields.

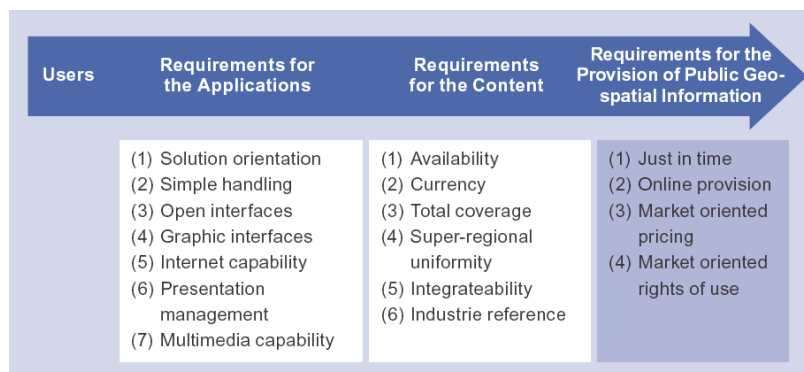
In the USA the University Consortium for Geographic Information Science (UCGIS) has identified a number of long-term research challenges in GIScience: (1) spatial data acquisition and integration; (2) interoperability of geographic information; (3) distributed and mobile computing; (4) future development of the spatial information infrastructure; (5) extensions to geographic representations (beyond two-dimensional, single-resolution maps); (6)

cognition of geographic information and ease-of-use issues; (7) scale; (8) uncertainty in geographic data and GIS-based analyses; (9) spatial analysis in a GIS environment; and (10) GIS and society (ethics, privacy); taken from Board on Earth Sciences and Resources (BESR, 2006).

The development of geographical information systems (GIS) has also made the demand for the automation of generalisation more evident. New tasks of information filtering and abstraction arise for instance in the context of spatial data mining and geographic information retrieval (Purves et al., 2007).

## 2.2 Production cartography

Traditionally paper maps, especially topographic maps are designed by professional cartographers with broad and general application scenarios in mind. Today, however, users increasingly expect to be able to process the cartographic data behind the maps, using their own GIS software or web mapping applications. Therefore mapping agencies have to provide digital data and services besides traditional paper maps. An analysis by Fornefeld et al. (2003) showed that the market of geoinformation in Europe was mainly supply driven, because the data providers offered products with available content, rather than considering the requirements of the user. Figure 2.3 illustrates the major requirements of the users. A distinction can be made between application requirements, requirements for the content and requirements for the provision for geospatial information.



**Figure 2.3**  
Major requirements of the users for applications, content and provision of geospatial information (Fornefeld et al., 2003).

Relatively new application requirements are the desire for multimedia capabilities, also available via web, accessible through simple and intuitive graphical user interfaces. The main requirement for the content refers to topicality and availability. Finally, on-demand access to geodata and geospatial information at a marginal cost is also important.

Meanwhile most European countries have recognised the value of geodata and react with the creation of national data infrastructures. In Switzerland as well a strategic concept was developed (GKG-KOGIS, 2003) leading to a package of measures dealing with basic geodata, meta information and geoservices, tariffs, guidelines for standards, technical infrastructure, education and further training as well as their legal foundation.

The reference data for the National Geodata Infrastructure (NGDI) in Switzerland will consist of the Topographic Landscape Model (TLM) with the corresponding Digital Terrain Model (DTM), the Digital Cartographic Model (DCM), orthophoto images as well as the data from the cadastral surveys. A main task for the National Mapping Agency (swisstopo) is the introduction of a GIS-based mapping production system, which supports the flexible derivation of cartographic models with automated and interactive methods.

**Figure 2.4**  
Planned derivation  
process of Digital  
Cartographic Models  
from Topographic  
Landscape Model at  
swisstopo (Feldmann  
and Kreiter, 2006).

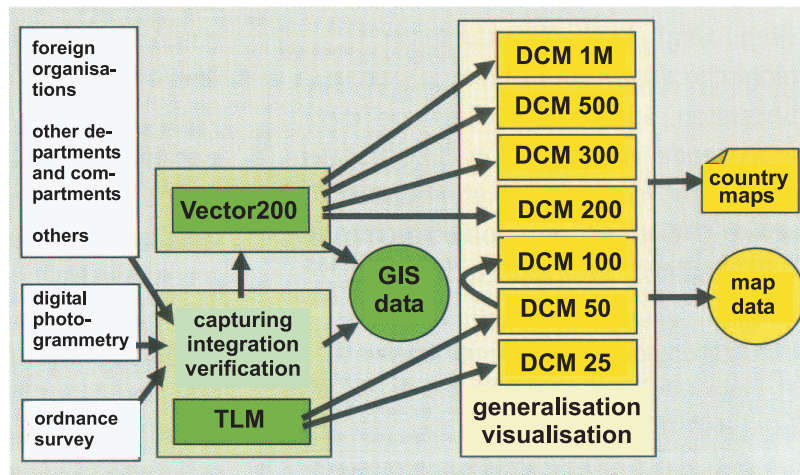


Figure 2.4 shows the data flow from the acquisition to the derivation of the final map series planned at swisstopo. The Topographic Landscape Model (TLM) forms the basis for the map pro-

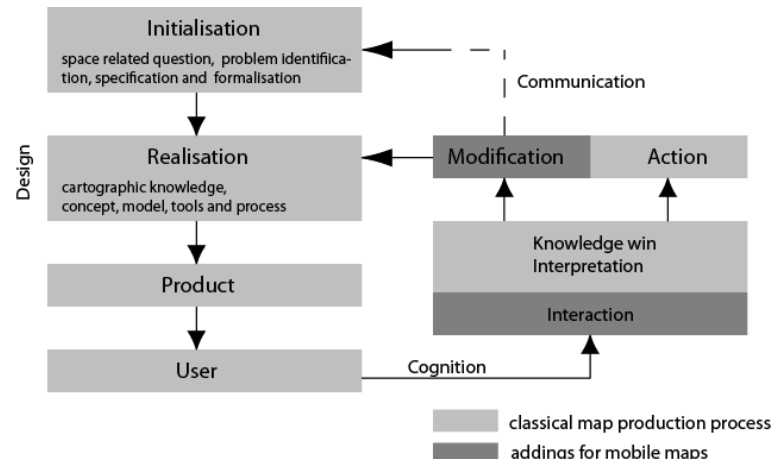
duction. The derived Digital Cartographic Models (DCM) should be connected with references and generalisation relevant meta-information to support further incremental updates. This functionality would optimise the complete production cycle and offer great development potential (Feldmann and Kreiter, 2006). Until now there exists no commercial GIS or map production system which has implemented this so called multi-representation database (MRDB) with explicitly modelled relations between features of different resolutions. Therefore a joint research and development project between the Department of Geography, University of Zurich and Axes System AG was carried out to extend the *axpand* map production system for the automated derivation of topographic maps from digital vector models (DRIVE - derivation of vector models), see Chapter 3.6. Sometimes a multiplicity of vector models already exists covering the same area, for instance the TLM and Vector200 (Fig. 2.4). In order to use and maintain these datasets effectively links have to be created independent from the generalisation process by means of automatic matching techniques (Lüscher and Burghardt, 2005).

The availability of freely selectable content by the user, independent from scale and extent has a vast impact on the mapping and especially the generalisation process (see Chapter 3). Minimal requirements for web mapping applications are proposed for instance by Origel-Gutiérrez (2004): title, place, current extreme coordinates, author, publication date, dynamic bar scale, frame or window for map contents, current legend, source of data and special semantic items for thematic documents. Kraak (2006) concludes that "not all maps might adhere to the cartographic design criteria, but let's not forget the 'fitness-for-use' motto". The classical map production process with map design, map original production and map reproduction is not suitable in the context of dynamic web mapping and mobile cartography.

As Figure 2.5 shows a more dynamic, iterative cycle, comparable with the "geographical inquiry model" (Malone et al., 2003) seems more appropriate. Scientific analysis and geographical inquiry involve a process of asking questions and looking for answers. The geographical inquiry model is made up of the following five steps: (1) acquire geographical resources (2) explore geographical data (3) analyse geographical information (4) act on geographical knowledge (5) ask geographical questions.

Map presentation and map interaction are essential for the geographical inquiry process. Kriz (2001) argues that in the past

**Figure 2.5**  
Iterative process of  
map creation, derived  
from [Kriz \(2001\)](#).



the map making process was lengthy, based on specialist work. Nowadays technology is available to offer simple as well as complex methods to the users for an interactive influence on the cartographic design process. The new tasks for the cartographic specialist include the design and development of this kind of methods and applications. The advantage of such systems is based on the fast realisation of the user's intentions.

### 2.3 Mobile cartography and location-based services

"The convergence of geospatial technologies - GIS, cartography, remote sensing, GPS, mobile computers, wireless communications and the Internet - has opened a new computing era." With these enthusiastic words from [Clarke \(2004\)](#) starts the foreword of a special issue of the journal *Cartography and Geographic Information Science* on "Mobile Mapping and Geographic Information Systems". What are the challenges in this new domain? Following a definition from [Reichenbacher \(2001\)](#):

*Mobile cartography*

*Mobile cartography* deals with theories and technologies of dynamic cartographic visualisation of spatial data and its interactive use on portable devices anywhere and anytime under special consideration of the actual context and user characteristics.



Adding to these main challenges are (1) the automated acquisition of user requirements, (2) the adaptation of map content and presentation to context and user characteristics and (3) the development of methods and hierarchical data structures for real-time geo-processing including generalisation. These three main topics will be investigated in this section more in detail.

The development of handheld devices enables users again to take digital maps out in the field, something which was already possible with paper maps - but impossible with interactive maps on desktop computers. In this sense [Scharlach and Müller \(2001\)](#) are right with their statement "the mobility of maps is nothing new". In their argumentation the term "mobile" is assigned exclusively to the media. In contrast other authors argue with new adaptation possibilities of mobile maps. Here, "mobile" refers to the fact that maps can react to the mobility of the user. Indeed, *interactive use* and *individual adaptation* are impossible with paper maps. Table 2.1 shows some characteristic differences between traditional maps and mobile cartographic presentations:

*Interactive use and individual adaptation*

characteristic	paper maps	mobile maps
interaction	static	interactive
map design	general	specific
content and map style	fixed	modifiable
target group	large community	individual user
hardware dependency	independent	dependent
map size	large	limited

**Table 2.1**  
*Characteristic differences between traditional maps and mobile cartographic presentations.*

A more detailed comparison is given by [Meng \(2005\)](#) with the characterisation of descriptive maps, analytical maps, exploratory maps, internet maps and mobile maps. Furthermore a distinction is made between *geocentric* and *egocentric maps*. While traditional geocentric maps are designed for a relatively long lifespan and diverse user groups, egocentric maps communicate geo-information for short-term usage by individual users. With these characteristics the egocentric map presentations as shown for instance in Figure 2.1 are much more appropriate for mobile usage scenarios. A similar presentation style is proposed by [Winter and Tomko \(2004\)](#) with a driver's metaphor realised through a shift of the current position of the map reader to the bottom of a mobile map.

*Geocentric and egocentric maps*

As mentioned in the definition of "Mobile cartography" the consideration of the actual user context is crucial for adaptation. [Dey \(2001\)](#) gives an overview of previous attempts to define and provide a characterisation of context and context-aware computing.

Context ([Dey, 2001](#))

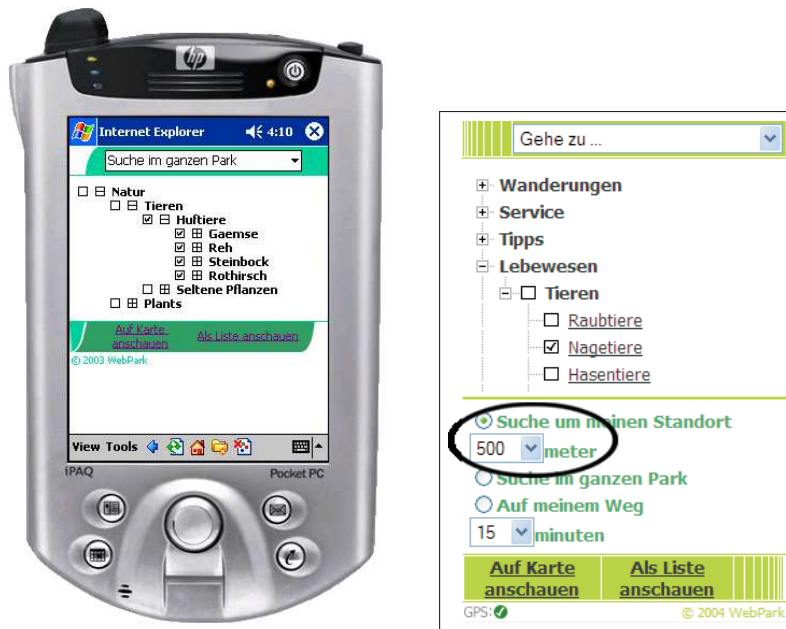
*Context* is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.

Context awareness in the domain of mobile cartography and mobile services is investigated and discussed for instance by [Nivala and Sarjakoski \(2003a,b\)](#) and [Weissenberg et al. \(2006\)](#).

**Automated acquisition of user requirements** As argued above new challenges in modern cartography refer to the consideration of individual user requirements regarding map content, abstraction level and presentation style. Necessary information about personal preference can be specified directly by the user or even better can be inferred automatically from user behaviour ([Fischer, 2001](#)). This so called dynamic profiling is based on the analysis of user activities and actions to determine what the user is interested in ([Weakliam et al., 2005](#); [Doyle et al., 2006](#)). A lot of research has been carried out in the human-computer interaction domain to capture user behaviour based on implicit feedback techniques, for an overview see [Kelly and Teevan \(2003\)](#) or [White \(2005\)](#).

*Action intention* A classification of the user's intention is made by [Chen et al. \(2002\)](#) into two levels: action intention and semantic intention. The *action intentions* are carried out on a lower level, such as mouse clicks, keyboard typing, following hyperlinks in a Web page, saving pages and the objects in a page or closing the browser window. Semantic intentions correspond to tasks at a higher level, potentially involving several basic actions. For instance "I want to see some marmots in the park" is a semantic intention typical of a visitor of an alpine nature park. [Chen et al. \(2002\)](#) introduced further the term of conceptual generalisation of keywords. Automatic extraction of keyword features delivers important meaning of sentences, but they are still too specific to represent the user's intention. Therefore generalisation of keywords into concepts can be carried out, through the usage of hierarchies consisting of synonyms (similar terms), hypernyms (broader terms), or hyponyms (narrower terms), etc. For instance, the hypernyms of the word





**Figure 2.6**  
Structuring of non-spatial information based on the generalisation of keywords according to [Chen et al. \(2002\)](#) in the Webpark mobile information system. User interface with geographical filter (black ellipse).

”marmot” in the hierarchy are marmot (*Marmota marmota*) → rodent (*Rodentia*) → mammal (*Mammalia*) → ... → animal → life form → entity. The concept hierarchy merges different animals on several levels such as mammal, animal, or even entity. The generalisation of keywords was exploited for the structuring and presentation of non-spatial information shown in Figure 2.6, developed in the context of the WebPark project (see also Section 3.6, and Chapters 9 and 10).

A prediction of user intention can be made if the generalised keywords are brought into relation with recorded user interactions. In the mobile environment the mapping of spatial data together with the user location receives a certain importance. Therefore additional types of user interactions can be gathered from user map interactions implicitly. Map browsing actions include panning, zooming, toggling features on/off, and highlighting map features ([Doyle et al., 2006](#); [Tezuka and Tanaka, 2005](#)).

On the level of *semantic intentions* as introduced by [Chen et al. \(2002\)](#) investigations were carried out by [Mountain \(2006\)](#) on the individual behaviour of tourists utilising mobile information systems. Geographical knowledge can be extracted from databases of users’ spatio-temporal behaviour to generate summaries and patterns to rank georeferenced information. The relevance of the information can be evaluated considering context information as the

*Semantic intention*

individuals' location and previously displayed spatial behaviour. As Mountain (2006) found out there are distinct classes of geographical context, which can be encapsulated by geographical filters such as: the "search around me filter", the "accessibility filter", the "search ahead filter" and the "visibility filter". The corresponding user interface for the first two geographical filters is shown in Figure 2.6 (black ellipse).

#### *Adaptive maps*

**Adaptive maps** A step further in the direction of considering the user requirements goes the idea of adaptive maps, which can be defined as a map or map-like visualisation adapted to a significant extent autonomously by a system or service Reichenbacher (2004). Adaptation allows the adjustment of content and map presentation, while required user interactions are minimised.










#### *Location-based services*

A very effective way of content adaptation is the consideration of the user location for information filtering, so called location-based services (LBS). A wide range of applications can be found for example in navigation, resource management, commerce, emergency calls, tourism, communication and entertainment. While adapting information to the user's location, the spatial scope over which the certain activities are performed can be considered as well. In Edwardes et al. (2003b) three different categories are proposed: a) immediate surroundings, b) region of activity or c) the whole resort. Zhang et al. (2003) recommended also to extend the spatially filtered information with a validity region around the client's location. This approach enables mobile clients to determine the validity of previous queries based on their current location.

An example for the adjustment of map presentation is the age specific adaptation of presentation style as shown in Figure 2.7.

In contrast to the abstract symbols for adult users, the symbols for children and teenagers are presented in a picture-like style. The icons for the older users are increased in size, assuming that older people have frequently difficulties with reading the small signs (Patalaviciute et al., 2005). Similar proposals were made earlier by Nivala and Sarjakoski (2004) as part of the GiMoDig project (<http://gimodig.fgi.fi>), with the adaptation of content and map presentation to the context of the user, the time of year and the purpose of use.

**Real-time generalisation and geo-enabling** A main task for offering location-based services is the preparation of spatial and non-

POI type	High school students + families	Adults interested in nature (18-50)	Adults interested in nature ( > 50)
Visitor center			
Picnic area			
Birdhouse			

**Figure 2.7**  
Examples of different  
symbols for different  
target groups  
(Patalaviciute et al.,  
2005).

spatial data. During the WebPark project (Chapters 9 and 10) it became obvious that data are not commonly found in a form that is fit for the purpose of an LBS. First of all, the data had originally been collected mostly with an alternative purpose in mind. Secondly, the tourism information and multimedia content analysed did not have an active geographical component. Therefore associations were created between spatial and non-spatial information items to make them useful for location-based service. This process was termed *geo-enabling*.

*Geo-enabling*

In the PhD thesis by Edwardes (2007) a number of research questions are investigated in detail related to the visualisation of geographical information in a mobile environment. An architecture for LBS is presented based on the Model-View-Controller design pattern covering the data component, portrayal and interactional respectively interface components. Furthermore the dynamic aspect of portraying geographic information is discussed in relation to user activities (Edwardes et al., 2003b). Finally a deformable grid structure is proposed to maintain the impression of spatial relationships both between the foreground map features and the base map, as well as absolute and relative positions of foreground features themselves (Edwardes et al., 2005c).

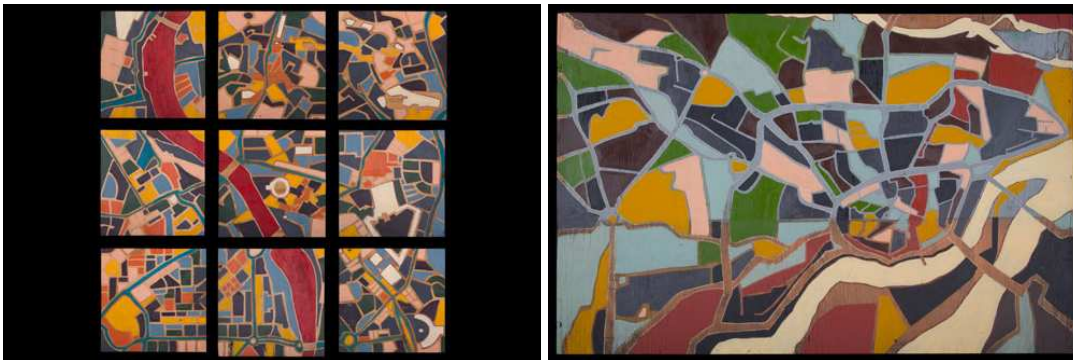
The generation of individualised cartographic presentations as a result of adaptation and personalisation requires geocoding in real-time, one of the big challenges for mobile GIS (Clarke, 2004). Solutions might be derived by an interplay of the three components client-server architecture, real-time methods and hierarchical data structures. The description of an architecture for the auto-

matic generalisation of mobile maps in the context of the WebPark project is presented in [Burghardt et al. \(2004a\)](#). A comparison of several architectures is made in [Edwardes et al. \(2003a\)](#) addressing the need to build an open generalisation system ([Edwardes et al., 2007](#)). The development of real-time methods is an upcoming research topic ([Lehto and Sarjakoski, 2005](#); [Sarjakoski et al., 2005](#); [Weibel and Burghardt, 2007](#)), selected approaches dealing with real-time map labelling for mobile applications as well ([Zhang and Harrie, 2006](#)). Data preparation also includes the creation of hierarchical data structures such as quadtrees, dendrograms or hierarchical stream ordering ([Burghardt et al., 2004b](#)). In Chapter 9 it will be shown how these data structures are exploited to facilitate real-time generalisation for the efficient presentation of thematic point data on portable devices.

## 3 Theory of automated map generalisation

### 3.1 Abstraction and generalisation

Generalisation can be described as the process of presenting general concepts by abstracting common properties of instances, e.g. see [Kramer \(2007\)](#). The ability of abstraction is an important condition for fast perceptive faculty and with it for effective intellectual grasp. In learning psychology the term "progressive abstraction" exists ([Kuehne et al., 2000](#)), which describes the capability of combining similar information by increasingly generic terms and as a result link-up knowledge in fine-meshed way.



In art the work is characterised as abstract, if it departs from figurative views. Sixlizards (<http://www.sixlizards.com>), a graphic design firm founded by Alice Dodge in 2002, presents pictures from maps in an artistic way by making them even more abstract than usual. Two pictures are shown in Figure 3.1. The artist Alice Dodge argues on her Web page "I hope to expand our thinking about representations of place - how we see, ignore, and experience our landscape."

**Figure 3.1**  
*Seville 2004, 9 panels each 32" x 32", oils on used plywood (left); Avignon 2004, 48" x 32", oils and oil pastel on used plywood (right) by Alice Dodge.*

Similar questions have to be investigated by scientists, when dealing with cartographic generalisation. [Mustière et al. \(1999, 2000\)](#) are subdividing cartographic generalisation into knowledge abstraction and representation. While knowledge abstraction tries to identify suitable geographic entities relevant to map purpose, the knowledge representation deals with the process of symbolisation. A standard definition from the International Cartographic Association ([ICA, 1973](#), p. 173) states:

*Cartographic  
generalisation*

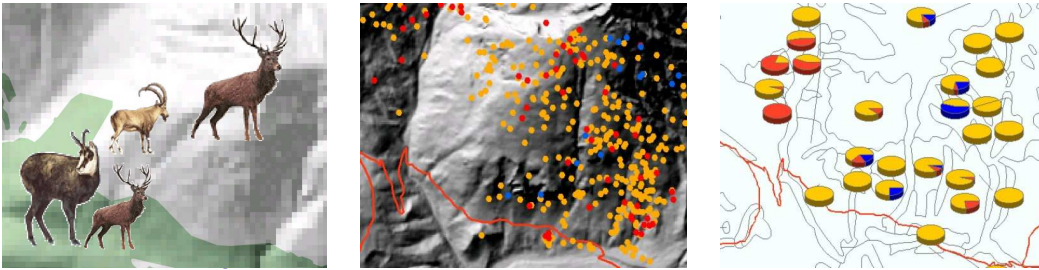
Cartographic generalisation is defined as "the selection and simplified representation of detail appropriate to the scale and/or purpose of a map".

One simple reason that maps differ in their level of abstraction is due to the differences in scale ([Peterson, 2006](#)). An interesting research problem refers to the question whether the level of abstraction can be operationalised in a way that the user can control it. A first example for the increase of the degree of abstraction in cartography exists with the selection of portrayal principle, for instance, with the change from picture-like presentations to symbolised maps. Even more abstract forms of portrayal include statistical representations or cartograms (see Section 3.2). Other prominent examples of abstraction in cartography exist with the usage of contour lines for the representation of a 3D phenomenon in two dimensions, or with the illustration of movement and migration flows through arrows.

Since the objective of mapping is to visually describe some of the characteristics of spatial phenomena, maps will always contain a certain degree of spatial and a certain degree of symbolic information. In [Edwardes et al. \(2005c\)](#) differences are highlighted between an abstraction of features and attributes (symbolic aspects) on one hand and an abstraction of spatial relationships (spatial aspects) on the other. Figure 3.2 shows an ordering of increasing abstraction for a selection of map views used in the WebPark project. Other examples are given by [Nivala and Sarjakoski \(2004\)](#) or [Patalaviciute et al. \(2005\)](#), see also Figure 2.7.

The first example in Fig. 3.2 attempts to describe data on animal observations by emphasising symbolically the attributes of each individual observation. It uses icons with a picture-like style, which are less abstract, intended to be suitable for children or teenagers (see also the Paragraph on *Adaptive maps* in Section 2.3).





**Figure 3.2**  
 Portrayal schemes  
 showing differing  
 degrees of spatial  
 relationships  
 (Edwardes et al.,  
 2005c).

The second example uses simple coloured dots to describe locations and number of animals observed. It emphasises the spatial aspects of the information such as the pattern of density and distribution. The degree of abstraction is increased through the usage of colour and the simplified shape. But the presentation style keeps still a link with reality, because the reader can imagine to see the animals from a far distance as point objects.

The third example describes information about the diversity in types of animals observed at the same location with help of a pie chart graphic. The presentation style is quite abstract giving place to express clearly the ratio of observed animal for the different species. The interpretation of the map might be more difficult, because the reader has to find out which relations are presented. On the other hand more detailed and interpreted information can be given.

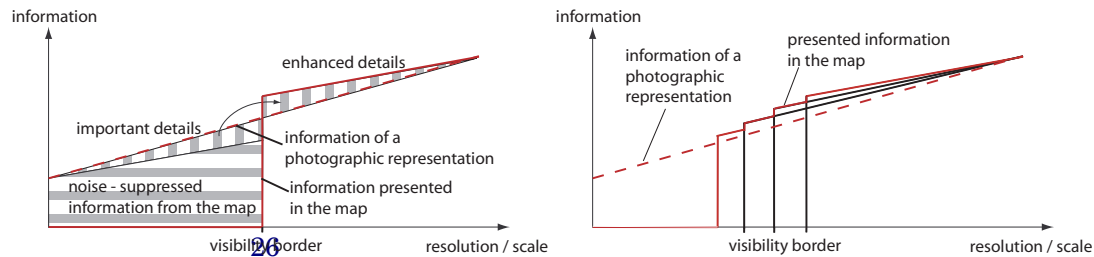
Depending on map purpose and user preference different portrayal schemes can be used to present information on different abstraction levels. While the first picture highlights the fact that there are animals, the last picture gives more detailed information about the animal distribution at different places. In this sense it seems possible to parameterise the degree of map abstraction giving a choice to the user for an appropriate presentation.

## 3.2 Continuous generalisation

**Discontinuities introduced by abstraction** The difference between maps and photographic presentations of the reality is based on the "essence" of abstraction (Board, 1967). This essence is brought into maps through the process of generalisation. This becomes obvious on the visibility border, represented by the so called minimal di-

mensions (SSC, 2005, p. 30). First, details or complete features are hidden, if their extent falls below the minimal dimensions. In this way the amount of information can be reduced, which can be interpreted as a filter operation removing the noise from the map. Second, exceptions are made if the suppressed details or features have a certain importance related to the purpose of the map.

Figure 3.3 tries to illustrate this process. As it can be seen some important information around the visibility border will be enhanced or enlarged, such that their extent becomes larger than the minimal dimensions. In contrast, less important information will be suppressed (i.e. removed from the map display). The consequence is an introduction of *discontinuities* visualised in Figure 3.3 through the sawtooth shape of the curve.



**Figure 3.3**  
*Discontinuity introduced by abstraction and generalisation.*

There are several generalisation operators such as simplification, elimination, collapse, aggregation and typification which remove or modify detail or complete map features in an abrupt way. These operations are carried out depending on the scale of the map and the corresponding minimal dimensions.

In traditional production and map use these discontinuities did not attract much attention. The reason is that the map series are produced in large scale intervals and the users have learned to reorientate while switching between two maps. With the introduction of web mapping applications the importance of discontinuities changed dramatically. Switching between map scales in real time became much easier and also the extent of the scale steps was reduced. Finally the user was given the possibility of free scale selection in the zooming process. At this point the need for continuous zoom functionalities became obvious.

Web mapping applications would ideally be based on maps that are generated "on-the-fly" at arbitrary levels of detail. Until now, all demonstrations of fully automated generalisation from a single



large scale database, however, are limited in the range of generalisation operations or they do not perform in a time that is applicable to online generalisation (Jones and Ware, 2005; Weibel and Burghardt, 2007). An alternative strategy is the utilisation of Multi-Representation Databases (MRDBs) with a limited number of pre-generalised maps.

**Smooth transitions for adaptive zooming** Following this, the question arises how the changes made to a map during generalisation can be visualised, such that the transition looks smooth and no sudden changes occur during the *adaptive zoom*. Adaptive zooming denotes the adjustment of a map, its contents and the symbolisation while changing the scale of a map (Cecconi and Galanda, 2002). A simple definition reduces adaptive zooming to the combination of map scaling with generalisation.

*Adaptive zoom*

Continuous versions of common cartographic generalisation operators were proposed in a conceptual paper by van Kreveld (2001). They are based around the basic operations:

- ❑ **moving** of features or feature details
- ❑ **morphing** between features or feature details
- ❑ **fading** into the background
- ❑ **appearing** from the background

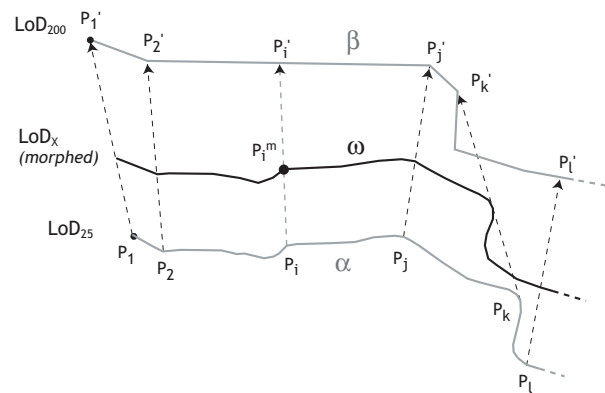
For example the *elimination* operation can be realised by shrinking a feature to a point through morphing or fading it into the background. Analogously, for *simplification*, *aggregation* and *collapse* the application of morphing operations yields smooth transitions. The *displacement* of a feature can be presented to the user in a continuous way through a stepwise movement of the displaced features. The *typification* operation is maybe the most difficult operation with respect to continuous transitions. Therefore a combination of moving, fading and appearing may be applied according to the proposal of van Kreveld (2001). As pointed out there, further research should address the organisation and development of intermediate representations of the data to allow smooth generalisation, as well as the question of how the different methods proposed can be integrated into a whole to obtain a pleasing overall result.

Sester and Brenner (2003; 2004a; 2004b; 2005) dealt with continuous generalisation of building ground plans. In order to achieve

this aim, they introduced a set of elementary generalisation operations (EGO's) such as *remove* and *extrude* and animated them. The animation is based on gradual shifts of the vertices similar to the moving operation by van Kreveld (2001) to reduce the popping effect of discrete changes. The usage of remove and extrude operations was inspired by elementary collapse and insertion operations of mesh simplification (Hoppe, 1996). The animation of such operations was suggested by Hoppe (1998) for the insertion or removal of nodes in a triangular network. The typification of settlement areas based on mesh simplification was proposed in Burghardt and Cecconi (2003) and will be described more in detail in Chapter 8.

Cecconi (2003) applied also morphing transformations for the interpolation of intermediate representations from different resolutions of a Multi-Representation Database (MRDB). The resolutions of the MRDB ( $LOD_{25}$  and  $LOD_{200}$ ) were derived from different sources, namely from 1 : 25 000 and 1 : 200 000 scale databases, respectively. Cecconi (2003) proposed a two-step procedure: 1) finding the correspondence between the features of the different resolutions and 2) the interpolation process. The interpolation is carried out by a morphing transformation based on the distance computation between corresponding vertices. A drawback of the method is that additional vertices have to be introduced for features of lower resolution.

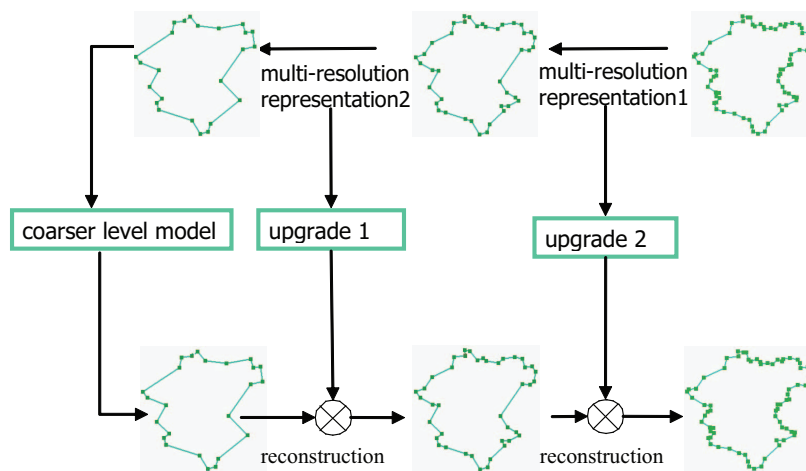
**Figure 3.4**  
The principle of a morphing transformation is to create an intermediate state out of two 'keyframes' represented by  $LOD_{25}$  and  $LOD_{200}$  (Cecconi, 2003).



**Progressive vector transmission for continuous zooming** Continuous zooming is best combined with progressive transmission of data from the server to the client. Much research has focussed on the rapid delivery of raster data, e.g. Dürst (1990) proposed the

Bitwise Condensed Quadtree (BCQ) which approximates an image in three dimensions, two spatial dimensions and the grey scale resolution, see also [Dürst and Kunii \(1991\)](#). Similar methods are used for an efficient compression and transmission of interlaced GIF or progressive JPEG images ([Burger and Burge, 2005](#)).

The delivery of vector data was investigated, for instance, by [van Oosterom \(1995\)](#), [Bertolotto and Egenhofer \(1999, 2001\)](#) or [Buttenfield \(2002\)](#). Successful implementations are described mainly for the transmission of triangle based vector data, see ([Hoppe, 1998](#)) and ([De Floriani et al., 2000](#)).



**Figure 3.5**  
The principle of progressive transmission of vector data ([Yang et al., 2007](#)).

The basic principle of progressive transmission of vector map data is illustrated in Figure 3.5 taken from [Yang et al. \(2007\)](#). Several key steps, namely multi-resolution representation, compression, and reconstruction are encompassed in the progressive transmission method. [Yang et al. \(2007\)](#) also present implementation details for the vector data delivery over the internet in terms of the preservation of consistent topology, transmission time and the cartographic quality of the resulting visualisations.

A interpolation method similar to the morphing approaches presented above makes it possible for mapping applications to offer continuous zooming, for example, from a continental view down to the street level, e.g. see <http://www.map24.com>. Every action appears smooth to the client and a complete reorientation as it is known from discrete scale changes is not necessary. The client application can also work asynchronously. It can start to display the data that is available on the client while it is wait-

ing to receive more data from the server. A hybrid mode of operation, including progressive transmission of vector and raster respectively image data is applied on the mapping application by <http://maps.google.com>. They offer this functionality also for mobile devices.

**Area cartograms and fish eye zoom** Continuous shape deformation methods are very important for dynamic mapping and mobile applications. Fundamentals are developed amongst others in the context of 3D visualisation (Sederberg and Scott, 1986; Hoppe, 1998) and pattern recognition (Kass et al., 1987). The term "homotopic" is used in mathematics to describe a continuous transformation between two mappings of the same topological space. Applications of homotopic transformations exist, for instance, in shape similarity comparison. Sako and Fujimura (2000) used homotopic deformations to capture global shape characteristics.

*Area cartogram*

There is also a longer tradition in cartography with continuous shape deformation for the creation of area cartograms (Tobler, 1973, 1986, 2004; Dorling, 1996; Keim et al., 2005). An area cartogram, also called value-by-area map, density equalising map or anamorphosis, is a map in which the sizes of regions appear proportional to related numerical information, while maintaining some degree of geographic accuracy. The maintenance of shape, topology and adjacency of regions are clearly valued in such maps. Different methods are developed for the derivation of area cartograms, for instance, pseudo-cartogram method (Tobler, 1986), line integral method (Gusein-Zade and Tikunov, 1993), constraint-based method (House and Kocmoud, 1998) or diffusion-based method (Gastner and Newman, 2004). An overview is given, for instance, by Tobler (2004) or Gastner and Newman (2004, 2005).

Possible mapping applications of cartograms are the representation of census results, election results, disease incidence, and many other kinds of human data. Some nice examples are shown with data from the Swiss Federal Statistical Office by Herzog (2003) (see also <http://www.mapresso.com/>) based on an implementation of algorithms by Dougenik et al. (1985) and Dorling (1996). There are also interesting applications of non-continuous cartograms, in which regions adjacent in reality are not adjacent on the cartogram, for instance, in Dorling's (1992) circular cartograms. Purves et al. (2005) used this method in the context of a spatially-aware information retrieval system, for the exploration of variation in document counts per region.

Besides continuous map deformation over scale and time, there are useful applications with continuous deformations in a single map. The method is called "fisheye zooming" and was originally developed by Furnas (1981, 1986). The main idea of the fisheye method is that objects in the centre are enlarged and hence presented clearly and distinguishable to the user, while objects which are further away from the centre are only shown at a reduced level of detail. The fisheye magnification creates an overview and maintains details at the same time. Frequent applications can be found within so called "Zoomable User Interfaces" (ZUIs), introduced by Donelson (1978) and further developed by Perlin and Fox (1993). In such user interfaces the fisheye graphical visualisation technique is applied to linear menus or icon bars like on the Apple Desktop with support of efficient searching (Bartram et al., 1995). A similar technique is called "liquid browsing", which uses an expansion lens based on a distance manipulation. In contrast to fish-eye zooming the information objects themselves are not enlarged, but rather the space between them (Waldeck and Balfanz, 2004).

*Fisheye zooming*

*Liquid browsing*

Further usages are imaginable in the context of mobile information systems (Harrie and Sarjakoski, 2002; Rappo et al., 2004). Typically the display size of the devices is very limited, which can partially be compensated for through user interactions such as zooming and panning. In the MSc thesis by Rappo (2003) several fisheye projections were investigated. Additionally a combined approach with cartographic generalisation is proposed to present more details in the centre of the magnifying glass and less information outside the zoomed area. The technical feasibility has been demonstrated and some results are illustrated in Fig. 3.6.

Fig. 3.6 shows the fisheye zoom with linear radial projection. The advantage of this projection is that all view lines from the centre of the map are straight in contrast to the orthogonal projection. In this example the fisheye zoom is applied to two test areas, an urban (left) and a rural one (right). At the bottom, both situations are generalised through a selection of streets and buildings. Additionally the settlement area is highlighted in the maps at the bottom. As a result more information can be shown (relatively speaking) in the centre of the zoomed area, typically the location of the user. As to most users, the appearance of fisheye projections will be a rather unfamiliar way of looking at maps, intensive user tests regarding the perception of this presentation style will have to be carried out, however.

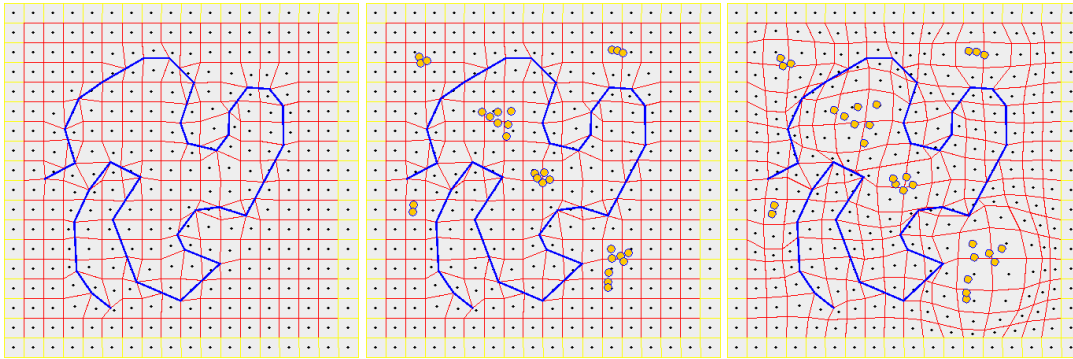
**Figure 3.6**  
*Fish eye projection  
 with linear radial  
 projection (urban -  
 left, rural - right).  
 The examples on the  
 bottom are  
 generalised through  
 the selection of  
 streets and buildings  
 (Rappo, 2003).*



A further development was carried out by [Edwardes \(2007\)](#) with a transformational grid approach applied to solve overlapping conflicts between dynamic map features, while preserving spatial and topological relationships amongst them. The basic idea is to separate general, static information necessary for orientational purposes as background from specific, dynamic foreground containing the requested information. Following that a grid can be aligned with the background information to create boundaries for the local transformations. Finally the foreground information is re-organised by a Laplacian smoothing transformation, which removes the overlapping conflicts. Figure 3.7 illustrates the approach in use for a small set of sample data.

The figure shows a re-organisation of points of interest using Laplacian smoothing. The primal grid is shown with red cells and the boundary cells, used to enforce the boundary conditions, shown in yellow. The small black points indicate the nodes of the dual





**Figure 3.7**  
Transformation by  
Laplacian smoothing  
(*Edwardes, 2007*).

grid (i.e the centroids). The grid has been initially fitted to a single feature line shown in blue. Foreground points are shown in orange.

**Perspective views** Transformed presentations, which are already accepted by the user are the perspective views well known from car navigation systems. A typical example is shown in Fig. 2.1 (left), taken from the TomTom Navigator (<http://www.tomtom.com>). Similar to the fisheye projection the information content varies with the distance from the imaginary position of the user. In contrast to fisheye zooming the presentation corresponds with the expected everyday real world view, which might be one reason for the acceptance of these visualisations.

The introduction of the third dimension into mapping applications brings the visualisation closer to the real world view and reduces normally the degree of abstraction. This leads to the question whether the level of abstraction might be a suitable control parameter for the user. For example current web mapping applications (e.g. <http://maps.google.com>, <http://maps.yahoo.com>, <http://local.live.com>) offer the possibility to switch between maps and orthophotos or look at them in a hybrid modus.

A simple evaluation related to 3D maps on mobile devices can be found in Coors et al. (2005). Their main conclusion is that the user's preferences and abilities are important factors in the context of selecting a presentation. Furthermore the task of the users and the map purpose determine which type of presentation might be useful.

### 3.3 Conceptual frameworks for automated generalisation

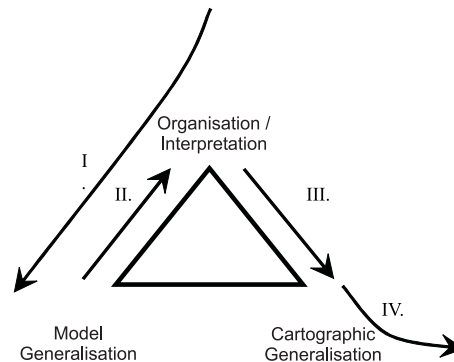
*Model generalisation*

*Cartographic generalisation*

*Semantic and geometric generalisation*

The generalisation process is often subdivided into model and cartographic generalisation (Müller et al., 1995). Model generalisation is responsible for adapting the map content to the required thematic and spatial resolution, e.g. by filtering, (re-)classifying and (re-)sampling data (Morgenstern and Schürer, 1999). This means that model generalisation determines which classes, objects and attributes are relevant at the required resolution of the target map or database. Cartographic generalisation, on the other hand, further considers symbolisation and graphical presentation of the map. It aims to provide a visualisation legible at the required cartographic scale and containing all the important information (Weibel and Dutton, 1998). Distinguishing between model and cartographic generalisation allows the treatment of content and presentation to be separated. In a similar way Jones and Ware (2005) separate between semantic and geometric generalisation. While semantic generalisation is related to the choice of the appropriate categories of information that should be represented, the geometric generalisation is concerned with the simplification of the shape and structure of the graphical symbols.

**Figure 3.8**  
Generalisation process of dynamically generated maps (Edwardes et al., 2003b).

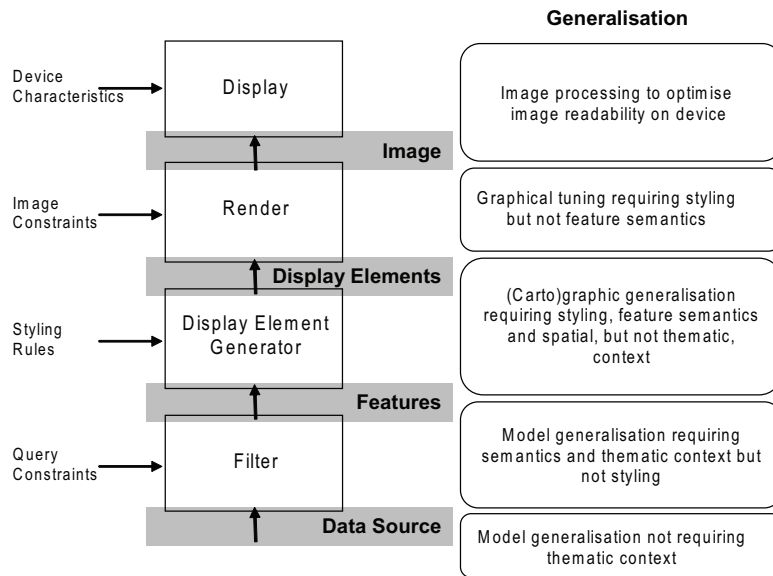


*Organisation and Interpretation*

In the context of the WebPark project a third component was introduced, with a focus on *organisation and interpretation*. The organisation and interpretation component serves the purpose of providing a map that is "fit-for-purpose", i.e. that is relevant to the questions that it is intended to answer. To this end the first step consist in organising the data into meaningful structures and groupings, e.g. patterns such as clusters and alignments which



should be communicated in the map. The second task is to provide design directions with the help of constraints, which determine what should be preserved and enhanced about these organisations in general (Edwardes et al., 2003b).



**Figure 3.9**  
The Portrayal model and generalisation model adapted from Cuthbert (1998) and following the generalisation typology of Weibel and Dutton (1998).

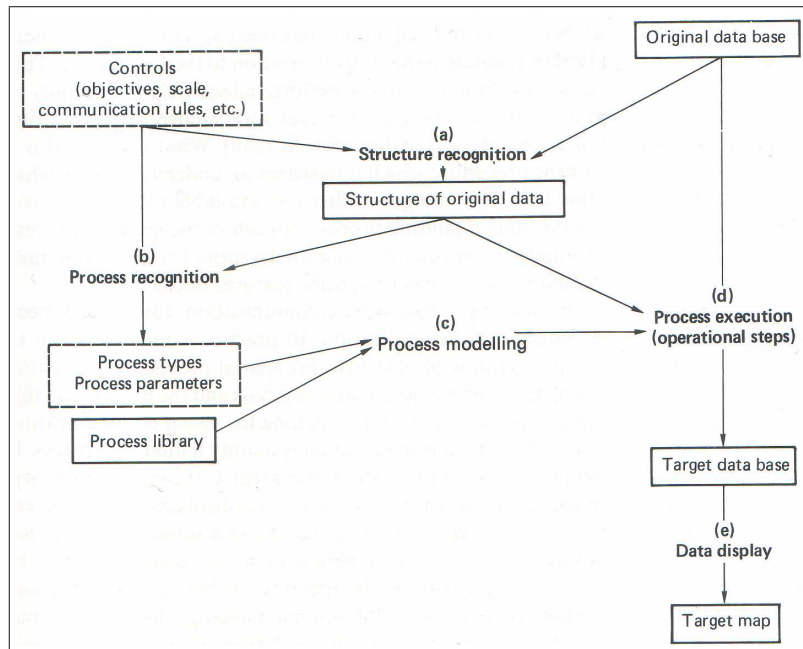
In the web mapping domain four stages of geospatial information are distinguished following Cuthbert (1998) – data holding (data source), accessing geographic objects (features), creation of graphical objects (display elements) and visualisation on the device (image). Several basic services are responsible for the transformation of the geospatial information. From data holding to map presentation the following processing steps need to be executed:

- ❑ **Filter:** Accessing the data by applying spatial and semantic filters, as well as provision of geometry and attribute information with the help of the Geography Markup Language (GML).
- ❑ **Display Elements:** Applying feature-centred rules of styling features e.g. with the Styled Layer Descriptor (SLD) mechanism for the creation of graphical objects.
- ❑ **Render:** Rendering of the map image with consideration of required image properties (size, number of colours, resolution).
- ❑ **Display:** Map respective image presentation on the screen.

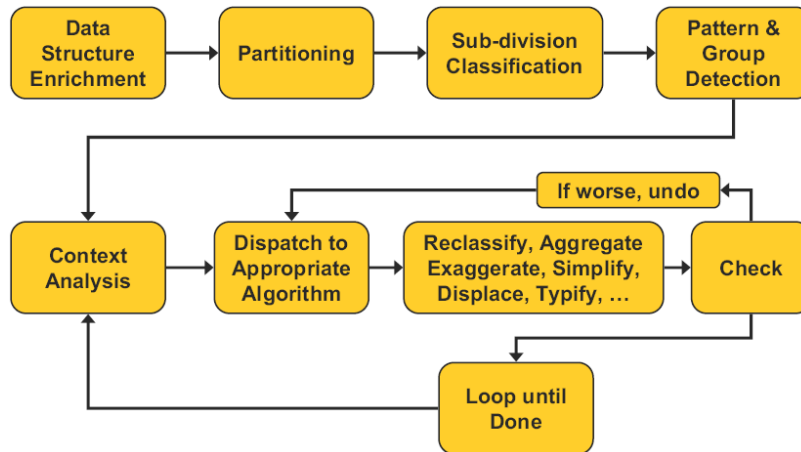
Generalisation functionality can be added to this model at various points. Fig. 3.9 illustrates this, presenting the extended portrayal model in Edwardes et al. (2007).

There exists a broad range of methodological frameworks for automated generalisation, summarised for instance by McMaster (1991) or Ruas (2002). An early conceptual framework was proposed by Brassel and Weibel (1988); it is shown in Fig. 3.10. Already at that time the generalisation process was subdivided into several tasks, including structure recognition, process recognition, process modelling, process execution, and the final map display.

**Figure 3.10**  
A conceptual  
framework of  
automated map  
generalisation  
(Brassel and Weibel,  
1988).



A detailed workflow model (Fig. 3.11) by Hardy and Lee (2005), which partially relies on earlier frameworks, made a distinction into pre-processing steps and a main generalisation loop. The pre-processing steps consist of data structure enrichment, partitioning, sub-division and classification as well as pattern and group detection. The main generalisation loop starts with context analysis, followed by the selection and execution of the appropriate algorithm and finishes with an evaluation. At the end a possible feedback loop directs either to the selection of another algorithm or a complete new context analysis.

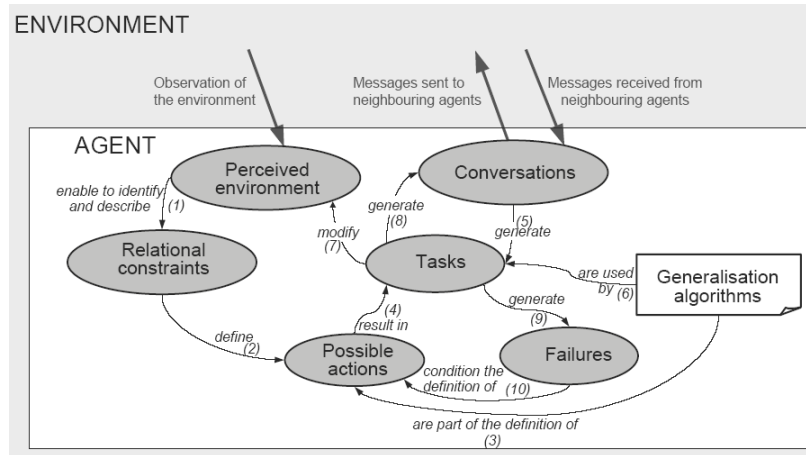


**Figure 3.11**  
Possible sequencing  
of generalisation tasks  
(Hardy and Lee,  
2005).

A quite successful framework derived from artificial intelligence was proposed by Ruas (1999) and applied in the European project AGENT (Barrault et al., 2001; Regnauld, 2001). There, features or groups of features were modelled through *micro*, *meso* and *macro* agents, which are able to identify and assess internal cartographic constraints. Based on constraint evaluation several plans of combined generalisation algorithms were carried out and tested against a cost function called severity. Hardy and Lee (2005, p. 186) argue, in advocating their workflow-based approach, "... a concept of individual features becoming self-aware software agents with a rigid lifecycle is unlikely to be applied, as experience by the primary author during and after the AGENT project indicates that this introduces unnecessary complexity and computational overheads for the achieved returns". Despite this comment, the idea of applying constraints for the description of an ideal cartographic situation was quite advanced, as well as the introduction of hierarchies through the categories of micro, meso and macro agents.

Fig 3.12 shows the conceptual model of an agent. The knowledge maintained by the agent about itself and its environment are organised in modules, depicted by grey ellipses. Arrows between ellipses represent the interactions between those modules. As the framework illustrates the approach is non-sequential, whereby the selection of generalisation algorithm is based on the (iterative) satisfaction of constraints.

**Figure 3.12**  
General architecture  
of an agent  
(Duchêne, 2004).

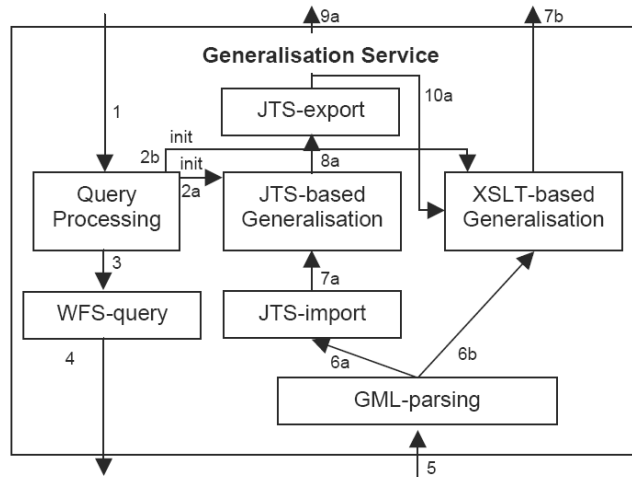


### 3.4 Technical frameworks for automated generalisation

In this section several technical generalisation frameworks will be discussed. An overview is given in [Edwardes et al. \(2007\)](#). Figure 3.13 describes the workflow of the generalisation service proposed in the EU project "GiMoDig – Geospatial info-Mobility service by real-time Data-integration and generalisation" ([Sarjakoski et al., 2005](#)). There are two computing environments for the generalisation service: XSLT and Java (JTS). [Lehto and Kilpeläinen \(2000, 2001\)](#); [Lehto and Sarjakoski \(2005\)](#) demonstrated the use of XSLT with Java extensions for real-time generalisation (e.g., selection and simplification).

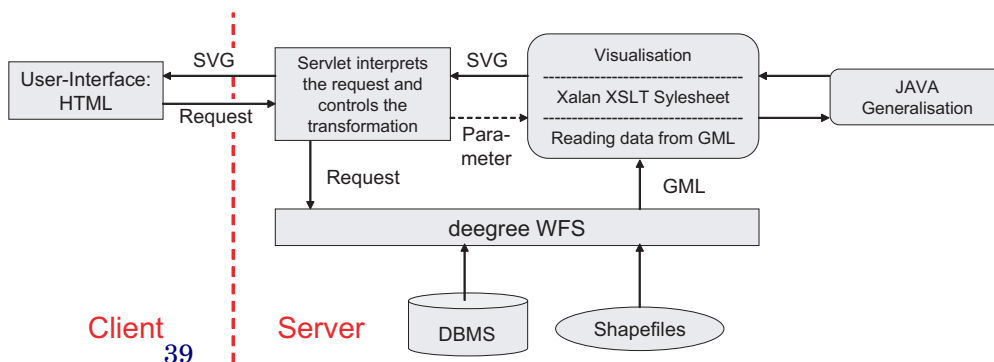
[Harrie and Johansson \(2003\)](#) and [Tuveson and Harrie \(2003\)](#) used the Java class libraries JTS and JCS from Vivid Solutions (2006) for implementing some real-time generalisation and data integration algorithms. The advantage of using XSLT is that it is computationally fast for *simple* generalisation tasks that treat each object individually. The Java environment based on JTS and JCS is not that computationally fast, on the other hand it provides tools for handling of *complex* relationships between objects in the generalisation process.

Our own developments showed similar results ([Mannes, 2004](#); [Burghardt et al., 2004a](#)). A framework based on XSLT, extended by Java classes, was developed for the real-time generalisation of point data representing animal observations. The corresponding



**Figure 3.13**  
Generalisation Service of GiMoDig, internal workflow. The numbers show the order in which the process steps are carried out. The branches a and b indicate alternative routes through the process (Lehto, 2003).

client/server architecture is shown in Figure 3.14. A number of hierarchical, partitioning and graph-based algorithms were implemented to aggregate the point data. Experiments showed that XSLT extended by Java classes is suitable for the aggregation of point data. However, in analogy to the statement made above, if more complex generalisation tasks have to be solved, a direct data access from Java is necessary.



**Figure 3.14**  
Client/server architecture for real-time generalisation based on XSLT extended by Java classes (Mannes, 2004).

The overview of Edwardes et al. (2007) gives an impression of the heterogeneity of generalisation frameworks inside the generalisation research community, yet it also testifies of the growing sophistication of the theoretical underpinnings of generalisation re-

search. Furthermore [Regnauld \(2006\)](#) deplores the fact that many researchers publish their algorithms, but the actual code or the executable is often short lived due to reasons such as researchers leaving academia, changes to the development platform, coding language becoming obsolete etc. This short life of algorithms leads to a limited interoperability with other data sets. There is a growing consensus in the research community that open research platforms would allow closer integration in terms of collaborative research, data abstractions, interoperability of functional components and the augmentation of geo-spatial applications with generalisation capabilities ([Edwardes et al., 2003a, 2007](#)). This desire has been evidenced through discussions at the various meetings of the ICA Commission on Generalisation and Multiple Representation (Beijing, 2001; Ottawa, 2002; Paris, 2003), described in [Edwardes et al. \(2003a\)](#).

A solution might be the application of a web service framework to overcome this interoperability problem. A prototype for the co-operation in the generalisation domain was developed in our research group and presented at the meeting of ICA Commission on Generalisation and Multiple Representation in La Coruña ([Neun and Burghardt, 2005](#)). A detailed description of the research platform based on web services is given in Chapter 4.

**Service framework for automated map production.** A joint project between the Department of Geography of the University of Zurich and Axes Systems AG is carried out to transfer scientific and research knowledge to practical applications. The project DRIVE / SerAx is promoted by the Commission for Technology and Innovation (CTI) Swiss Federal Office for Professional Education and Technology (OPET). The motto of KTI is "Science to Market". The aim is the rapid conversion of state-of-the-art laboratory findings to marketable products. With this goal the KTI promotes research and development projects between universities and enterprises.

The emphasis of the project lies in the extension of the *axpand system* (a software product of Axes Systems AG), in order to be able to accomplish the necessary automated steps for the production of topographic maps from digital vector models, including model and cartographic generalisation, in an integrated development environment. Research-relevant questions focus first of all on establishing links between digital vector models of different scales using multi-representation databases (MRDB), as well as the utilisation

tion of these linkages for the automatic update. A goal is to reduce the workload for data and map updates across related scales.

A second main focus concerns the research and development of generalisation operators, deep examination of these operators under consideration of "neighborhood" relationships, as well as the use of suitable data structures for contextual generalisation algorithms/operations. In addition, the parameterisation and controlling of the generalisation operators are to be improved. The goal is to replace static sequential batch methods of operation by dynamic workflows. The sequence of generalisation operators used should be specified and defined in a situation-dependent way. This should be achieved by shifting the selection of generalisation operators from the level of the object class to the level of the individual object. This requires a preceding automatic conflict analysis.

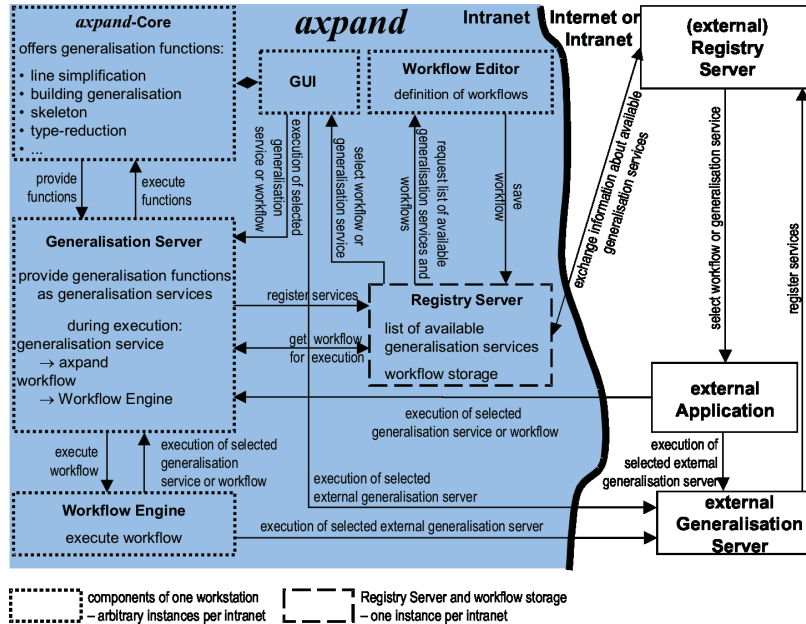
The service framework as shown in Figure 3.15 is composed of three main components: the core axpand system (axpand-Core), a generalisation service architecture (Generalisation and Registry Server), and a workflow management system (Workflow Editor and Engine). The generalisation functions of the cartographic GIS axpand were extended during the project DRIVE (Burghardt et al., 2005a; Bobzien et al., 2006a). The generalisation functions are offered as generalisation services via a so-called generalisation server. A central server, the so-called registry server, describes which generalisation service is available and where it can be accessed. Several generalisation servers could exist, offering a variety of generalisation services.

The system design considers two network scenarios: intranet and internet. The local, autonomous intranet is strictly separated from the internet. The internet scenario allows the access of external generalisation services and the usage of internal services by external applications. Additionally, a workflow management system will be integrated consisting of a workflow editor and a workflow engine. The workflow editor enables a graphical design of the generalisation workflows. The editor is provided with all possible generalisation services and workflows by the registry server. The workflow engine is part of the generalisation server. So, the generalisation server can execute workflows as well as single generalisation services.

The developed service platform will support also research on processing strategies of services in general and automated generalisation in particular, for instance the chaining of services and generalisation operators, the orchestration and parameter selec-



**Figure 3.15**  
Service framework of  
the commercial  
cartographic GIS  
expand (Axes  
Systems AG) with  
generalisation and  
registry server and  
the workflow  
management system  
(Petzold et al.,  
2006b).



tion as well as the determination of independent generalisation tasks. Furthermore important questions relate to the semantic and syntactic service description as a prerequisite for automated processing (Lemmens et al., 2006). This leads into the direction of sharing geo-processes, which was pointed out, for instance, by Goodchild (2005). Most of the research and the implementation studies on standards are devoted to facilitate the sharing of geographic data, but very little effort has gone into making it possible to share process objects, or representations of processes in the GIS and cartography domain. Chapter 4 will deliver an example of work that enables sharing process objects, in particular generalisation services.

### 3.5 Generalisation state-of-the-art

The previous section of this chapter have reviewed the recent developments in automated map generalisation and related topics over the past decade. In this section, there will be a start off by a historical overview of the roots of generalisation research and then work towards defining current research challenges through an analysis of "hot topics" in recent research workshops and journal issues.



### 3.5.1 Historical overview

The following section gives a short historical overview of research on automated generalisation. The main sources for this historical literature study are [Hake \(1985\)](#), [Brassel and Weibel \(1988\)](#), [Buttenfield and McMaster \(1991\)](#), [Müller et al. \(1995\)](#), [Meng \(1997\)](#) and [Weibel \(1997\)](#).

**The beginnings - the 1960ties and 1970ties** First theoretical investigations on the topic of automated generalisation were carried out in the middle of the 1960ties. Impressive approaches, partially usable until recent times are "experiments in the computer generalisation of maps" from [Tobler \(1964, 1966\)](#), " $\epsilon$ -filtering" by [Perkal \(1966\)](#) or the "radical law" by [Töpfer and Pillewitzer \(1966\)](#). An early conceptual framework of generalisation was proposed by [Ratajski \(1967\)](#) with the distinction of quantitative and qualitative generalisation. These pioneers established the foundations for further research on automated generalisation.

There were also older approaches from other disciplines which still receive attention in the generalisation domain such as the laws of organisation in perceptual forms ([Wertheimer, 1938](#)) or approaches for the selection in river networks based on stream ordering schemes ([Horton, 1945](#); [Strahler, 1952](#); [Shreve, 1966](#)).

Research and development at the beginning of the 1970ties were concentrated on the development of geometric algorithms, especially for the purpose of line simplification e.g. [Lang \(1969\)](#), [Boyle \(1970\)](#), [Ramer \(1972\)](#), [Douglas and Peucker \(1973\)](#), [Duda and Hart \(1973\)](#) and [Reumann and Witkam \(1974\)](#).

Further research mainly in Europe dealt also with building generalisation e.g. [Berger \(1974\)](#), [Staufenbiel \(1974\)](#), [Hake and Hoffmeister \(1978\)](#); displacement e.g. [Lichtner \(1977\)](#), [Christ \(1979\)](#) and [Schittenhelm \(1980\)](#); generalisation of object groups e.g. [Gottschalk \(1974\)](#), [Christ \(1977\)](#) or [Johannsen \(1977\)](#); as well as generalisation of road networks by [Menke \(1982\)](#).

First approaches which utilised raster data for generalisation of settlement areas were carried out by [Gottschalk \(1973\)](#). At the end of the 1970ties [Lichtner \(1979\)](#) discussed for the first time the problem of sequencing generalisation algorithms.

**The 1980ties and 1990ties** In the 1980ties researchers started with classifying available generalisation algorithms into typologies of generalisation operators, e.g. [Brassel \(1985\)](#), [Nickerson and Freeman \(1986\)](#), [De Lucia and Black \(1987\)](#) and [McMaster \(1989\)](#).

An often used typology was proposed by [McMaster and Shea \(1992\)](#) with a distinction of operators for attribute transformation (classification, symbolisation) and operators for spatial transformation (simplification, smoothing, aggregation, amalgamation, merging, collapse refinement, exaggeration, enhancement, displacement).

Early software systems used batch processes to sequentially chain together generalisation algorithms ([De Lucia and Black, 1987](#)). The most prominent example is the CHANGE system from the University of Hannover for the generalisation of buildings and streets ([Hoffmeister, 1978](#); [Powitz, 1988](#); [Grünreich, 1993](#); [Bobrich, 2001](#)). Improvements over the simple sequential approaches were reached with the application of rule-based or expert systems in map generalisation ([Mackaness et al., 1986](#); [Nickerson and Freeman, 1986](#); [Jankowski and Nyerges, 1989](#); [Robinson and Zaltash, 1989](#); [Weibel, 1991](#); [Buttenfield et al., 1991](#); [Schylberg, 1993](#)). The underlying methodology is termed condition-action model, whereby the generalisation process is subdivided into a structure recognition phase and an execution phase of generalisation operations. The rules of the expert systems can be interpreted as switches between these two phases, responsible for the condition dependent execution of generalisation actions (or operators).

To acquire the procedural knowledge for the expert systems a lot of research was spent on the topic of "knowledge acquisition" ([Weibel, 1993, 1995](#); [McMaster, 1995](#)). Various methods were utilised such as interviews with cartographic specialist ([Richardson, 1989](#)), interaction logging of cartographic system users ([Weibel et al., 1995](#)), extracting knowledge from handbooks of cartographic conventions and reverse engineering ([Leitner and Buttenfield, 1995](#)) as well as automated learning methods ([Werschlein and Weibel, 1994](#)). One of the major drawbacks of expert systems is the increasing number of necessary rules, which partially work against each other.

Therefore less restrictive frameworks were favoured as proposed for instance by [Brassel and Weibel \(1988\)](#) (see also Fig. 3.10). In this framework the generalisation process is decomposed into a) structure recognition, b) process recognition, c) process modelling, d) process execution, and e) data display. If all the phases are realised in an automated way the generalisation process execution will be much more flexible.

The framework can be seen as first step in the direction of constraint-based models. The underlying concept changed from modelling of the generalisation process through static path de-

scriptions, to a model where only start and endpoint are described by objectives (termed constraints), while the route in between is flexible. Constraints as introduced by [Beard \(1991\)](#) allow an evaluation whether the map is cartographically satisfactory or at least whether the cartographic presentation improves during the generalisation process.

During the 1990ties the diffusion of Geographic Information Systems (GIS) influenced cartographic map production and automated generalisation. Tool sets of interactive generalisation algorithms were offered such as the Map Generalizer from Intergraph or the ArcGIS toolbox from ESRI. The idea was that these generalisation tools should solve the simple routine work, while the cognitive work load of situation analysis and process control had to be carried out interactively by cartographic specialists.

A lot of research was invested in the application of auxiliary data structures in generalisation, for example graph theory ([Mackaness and Beard, 1993](#)), multi-scale trees ([Frank and Timpf, 1994](#)), hierarchical data structures ([van Oosterom, 1995](#)), Delaunay triangulations ([Jones et al., 1995](#)), quadrees ([Dutton, 1998](#)) and Voronoi diagrams ([Gold and Thibault, 2001](#)). Complete new research topics, motivated through the spreading of GIS evolved with model-oriented generalisation ([Morgenstern and Schürer, 1999](#)) and multi-resolution databases, e.g. [Egenhofer et al. \(1994\)](#), [Jones et al. \(1996\)](#), [Weibel and Dutton \(1999\)](#), [Spaccapietra et al. \(2000\)](#).

Since the end of the 1990ties the orchestration of the generalisation process attracts a lot of attention. The main tasks refer to the situation dependent selection of generalisation operator sequences with appropriate parameter settings. Approaches of two major categories were proposed, first the so called optimisation methods, which perform different operations simultaneously and second the constraint based approaches that perform one operation at a time ([Mustiere, 2005](#)). A similar categorisation was proposed by [Harrie and Weibel \(2007\)](#), which distinguishes between combinatorial optimisation modelling, continuous optimisation modelling and agent modelling. They argue that all three methods are based on constraints, despite the fact that the constraints are part of the objective function in the case of optimisation modelling.

The combinatorial optimisation techniques have been proposed in cartography first to solve the problem of map label placement by [Zoraster \(1986, 1997\)](#). Integer programming was used to formulate line simplification as a combinatorial optimisation problem ([Crom-](#)

ley and Campbell, 1992) and later genetic algorithms were applied to solve spatial conflicts (Wilson et al., 2003). Worth mentioning is also the approach of Ware et al. (2003) who use simulated annealing both for the processing of a single generalisation operator as well as the selection of generalisation operators.

Continuous optimisation methods range from spring models (Bobrich, 1996), snakes (Burghardt and Meier, 1997) and elastic beams (Bader, 2001), finite elements (Højholt, 2000) to least square adjustment (Sester, 2000; Harrie, 2001). Irrespective of the advantages of continuous methods as pointed out in Section 3.2, there are limitations with discrete operations such as aggregation, selection and typification.

Agent modelling in combination with object oriented methods became very popular through the AGENT project (Ruas, 1999; Regnauld, 2001), see also Fig. 3.12. There, the selection of generalisation operators was based on the satisfaction of cartographic constraints, which were summed up in a cost function called severity. Agent modelling is capable to handle the different types of generalisation operations. Furthermore also the optimisation techniques can be integrated, as demonstrated by Galanda and Weibel (2003). Duchêne (2003, 2004) extended the agent model with capacities to perceive their spatial environment, as well as an ability to communicate with surrounding agents. The AGENT model can be seen as the most important achievement of generalisation research at this time with a strong impact also on current research.

### 3.5.2 Generalisation workshops

A good overview of past and current research topics can be found in the publications of the *Generalisation Workshops* organised by the ICA Commission on Generalisation and Multiple Representation (<http://aci.ign.fr/>). Papers and presentations are usually grouped together in sessions. Table 3.1 categorises these topics of this sessions and reflects the main research interests of the past decade. Six main categories are distinguished.

The first category is related to NMA and vendor perspectives, containing papers describing both requirements as well as solutions from a production point of view. It is very encouraging to have this practical link at the beginning of the workshop, which opens brackets with the presentation of unsolved problems to the integration of scientific approaches in commercial products and their application for map production.

The second category covers aspects of quality assessment, knowledge formalisation and acquisition as well as approaches for feature conflict detection. These research topics were quite popular in the middle of the 1990s with the attempt to replace batch processing by expert systems and constraint based approaches. It seems that a new round of generalisation research was opened at that time reflected, for instance, by the beginning of this workshop series on automated generalisation. Therefore, it was quite natural to start with research on conflict detection to determine situations where automated generalisation was needed. Furthermore the research on quality assessment and knowledge formalisation was carried out to define satisfactory or ideal map solutions respectively, which should be reached later with automated solutions.

The third category of research is dedicated to the modelling of spatial and hierarchical relations for generalisation purposes. Most important are the topological relations which have to be preserved by the generalisation algorithms. Furthermore relations are needed to describe patterns such as alignments, neighbourhood relations such as a train station connected with a railway, or partitions such as buildings inside a street mesh. Hence, of the large number of possible spatial and semantic relationships, only a selection is usually modelled explicitly and stored persistently in a database. Most of them are more usefully computed *ad hoc* and on demand. In contrast semantic properties and relations needed for generalisation that cannot be derived from geometry must be coded explicitly.

The fourth category covers approaches and implementations of generalisation operators mainly devoted to line and area generalisation. Many investigations are made on line simplification and smoothing, but there is also research on the more elaborate generalisation operations such as displacement, aggregation or typification. Further improvements of generalisation operators will depend strongly on the research and developments of formalised relationships between map features as described in the third category above.

The fifth category contains research on Multi-Representation Databases (MRDB) with integrated or connected feature representations for different scales, resolutions and temporal states. Therefore this research category encompasses contributions on model generalisation as well as incremental update. Additionally research covers methods on the integration of feature representations of the same real world phenomena from different data sources

on the schema and instance level. While the schema integration investigates semantical issues, the integration of features based on matching methods has also to consider geometrical and topological aspects.

The sixth category finally combines work on the overall generalisation process. There are methods proposed which deal with the orchestration and combination of generalisation operators, as well as their situation dependent parameterisation. Representative approaches are on the one hand the optimisation methods with the simultaneous application of different generalisation operations and on the other hand the constraint based methods including AGENT which searches for suitable operator sequences.

### 3.5.3 Current research challenges

In Table 3.2 one further category of workshop contributions is presented covering recent and future research topics. Starting from that, possible research challenges are pointed out in the broader context of automated generalisation.

A first main topic is on-the-fly generalisation, which describes the application of automated generalisation in real-time. Until the end of the nineties, cartography was mainly devoted to static paper map products. With the spread of web mapping and mobile applications new possibilities of user interaction required flexible and fast presentation and generalisation methods to satisfy individual information requests. In this sense on-the-fly generalisation can also be seen as being a sub-set of the more general trend towards on-demand mapping and map adaptation.

In order to realise on-the-fly approaches two different strategies can be applied, either through the usage of fast and efficient algorithms in real-time or the utilisation of precalculated hierarchical (reactive, auxiliary) data structures. Real-time methods are more flexible, because their application is not restricted to given scales. On the other hand they are technically demanding and computationally more costly than hierarchical data structures.

Hierarchical data structures were successfully applied mainly for automated label or icon placement. Despite the fact that the requirements of on-the-fly generalisation are well known until now there is only relatively little research work being pursued on this topic.

The second category of upcoming research relates to 3D generalisation. Major research questions in this field dealing with the visualisation of large 3D city models and the modelling of 3D topol-

	Session topic	'95	'97	'99	'01	'02	'03	'04	'05	'06
I	production issues, NMA and vendor perspectives		X	X	X		X	X	X	X
II	quality assessment, measures and constraints	X	X	X	X	X	X			
	knowledge acquisition and encoding	X	X				X			
	feature conflict detection, shape analysis	X	X			X	X		X	
III	modelling spatial and hierachical structures		X	X		X				
	modelling semantics and non-spatial structures		X	X						X
	data base enrichment						X	X	X	
IV	feature conflict resolution, generalisation algorithms	X	X			X		X	X	
	line generalisation, smoothing			X	X	X				
	area generalisation				X					
	displacement					X				
	clustering, typification, amalgamation						X			
V	MRDB, database updating			X	X	X		X		
	model generalisation			X			X			
	matching								X	
VI	optimisation methods			X	X		X			
	general. process, agents, knowledge based methods						X		X	

ogy. The fast visualisation requires the representation of buildings at different levels of detail (LoD). The derivation of these LoD should be realised on the basis of automated generalisation procedures.

**Table 3.1**  
Session topics of  
Generalisation  
Workshops organised  
by the ICA  
Commission on  
Generalisation and  
Multiple  
Representation.



	Session topic	'95	'97	'99	'01	'02	'03	'04	'05	'06
1	real-time generalisation				X					
2	3D generalisation						X	X		
3	generalisation for mobile applications							X		
4	web generalisation services, platform, architecture								X	X

**Table 3.2**

*Sessions on upcoming topics.*

The third category of Table 3.2 describes generalisation research for mobile applications. Besides the requirement of real-time processing as pointed out above, two further constraints have to be considered. Mobile devices offer significantly less processing power and the screen size for information visualisation is rather limited. Consequently the selection, filtering and adaptation of information receives overriding importance.

The last category deals with the utilisation of web services for generalisation purposes. Services can be integrated into any software platform, such as web browsers, GIS, or map production software and they can be accessed either over the internet or locally. Furthermore they allow the interoperable use with different programming languages independent from the operating system and the hardware used. With that they extend the traditional approach of software libraries, which can often only be used in one programming language. Advantages of a service oriented architecture are also the capabilities to optimise processing capacities both between different servers (parallel processing) and also between server and thin or thick clients.

Research questions have to investigate, for instance, the degree to which a subdivision of services might be useful, limited through the network traffic generated and which kind of information should be stored persistently vs. repeated recalculation. Also of great importance are the development of mechanisms to create formalised, machine readable descriptions of service functionalities, as well as the standardised exchange of considered constraints and relations. Finally research has to identify further, alternative paths for generalisation service exploitation.



### 3.5.4 Selection of books and special issues on generalisation

#### Books on generalisation

- ❑ Buttenfield, B. P. and McMaster, R. B. (1991). Map Generalization: Making Rules for Knowledge Representation. *Longman Group*. ([Buttenfield et al., 1991](#)).
- ❑ McMaster, R. and Shea, S. (1992). Generalization in digital cartography. *Washington D.C: Association of American Geographers*. ([McMaster and Shea, 1992](#))
- ❑ Müller, J.-C., Weibel, R. and Lagrange, J.-P. (1995). GIS and generalization - Methodology and Practice. *Taylor & Francis, GISDATA 1* ([Müller et al., 1995](#))
- ❑ Kreveld, M. van , Nievergelt, J., Roos, T. and Wiedmayer, P.(1997). Algorithmic Foundations of Geographic Information Systems. *Number 1340 in LNCS. Springer Verlag*. with *Chapter 5*, pp. 99-152, from Weibel, R. (1997). Generalization of Spatial Data: Principles and Selected Algorithms. ([Weibel, 1997](#))
- ❑ Ruas, A., Mackaness, W.A. and Kilpeläinen, T. (2007). Challenges in the Portrayal of Geographic Information: Issues of Generalisation and Multi Scale Representation. *Elsevier Science*. ([Ruas et al., 2007](#))

#### Non-English books on generalisation

- ❑ Töpfer, F. (1974). Kartographische Generalisierung (German). *VEB Hermann Haack, Geographisch-Kartographische Anstalt Gotha / Leipzig*. ([Töpfer, 1974](#))
- ❑ Hake, G., Grünreich, D. and Meng, L. (2002). Kartographie - Visualisierung raum-zeitlicher Informationen (German). *de Gruyter Lehrbuch, 8. Auflage*. ([Hake et al., 2002](#))
- ❑ Ruas, A. (2002). Généralisation et représentation multiple. *Hermès Science Publications*. ([Ruas, 2002](#))

### Special issues with topics related to automated generalisation

- ❑ McMaster, R. (1989). Numerical Generalization in Cartography. Special Issue of *Cartographica*, Vol. 26, No. 1.
- ❑ Weibel, R. (1995). Map Generalization. Special Issue of *Cartography and Geographic Information Systems*, Vol. 22, No.4.
- ❑ Weibel, R. and Jones, C. B. (1998). Issue on Map Generalization. Special Issue of *Geoinformatica*, Vol. 2, No.4.
- ❑ Richardson, D.E., and Mackaness, W. (1999). Computational Methods for Map Generalization. Special Issue of *Cartography and Geographic Information Science*, Vol. 26, No. 1.
- ❑ Jones, C. B. and Mark, J. W. (2005). Special Issue: Generalisation. *International Journal of Geographical Information Science*, Vol. 19, No. 8-9.

**Commercial generalisation software and tools** The number of commercially available software packages for cartographic generalisation are rather limited, e.g. FME (Safe Software), ArcGIS (ESRI), Clarity (Laser-Scan), DynaGen (Intergraph), axpand (Axes-Systems-AG). An extended comparative study of generalisation software and tools was made in 1999 (Ruas, 2001). Besides the detailed results about the capabilities of the generalisation software tools, also some general research and development tasks were pointed out there:

- ❑ design and develop better descriptions of the data content (spatial analysis)
- ❑ develop conflict detection tools with respect to user needs and the property of geographical data
- ❑ develop more contextual algorithms such as object contextual removal (according to flexible criteria) and object displacement
- ❑ automate processes by learning techniques based on character and conflicts and experimental test

Further investigations were carried out in 2005 (Stoter, 2005) with focus on current generalisation practice in National Mapping Agencies (NMAs) from Europe. The title "Generalisation: the gap between research and practice" highlights some of the outcome, that

NMAs and software vendors have their responsibility to become aware of research developments and to report about their experiences. Similar observations were already made by Müller et al. (1995):

- ❑ Research cooperation between NMAs and academic research should be intensified. NMAs should state their requirements with respect to generalisation functions more clearly and academic research should take up these issues.
- ❑ Likewise, the third player in R&D, software vendors, should be in close contact with developments taking place at NMAs and sponsor research at academic institutions.

Currently, with the kick off meeting in October, 2006 at ITC (Netherlands), a EuroSDR research project on the state-of-the-art of generalisation is being carried out. The objectives can be summarised as "to establish by a small set of controlled tests the state-of-the-art in generalisation, particularly with reference to 'framework data' for map production, so as to inform both potential users and ongoing research" (Stoter et al., 2006). The project team consisting of members from both NMAs (IGN/France, OS Great Britain, KMS/Denmark, TDK/The Netherlands, ICC/Catalonia) and research institutes (University of Hannover, University of Zurich, ITC). Six software suppliers indicated their willingness to have their software tested in this project (LaserScan (Clarity), Intergraph (Geomedia), ESRI (ArcGIS), Axes, CHANGE/PUSH/TYPIFY (University of Hannover), CPA). The next challenges are the definition of the precise context for the tests, i.e. data sets, test procedure and evaluation procedure, in order to be able to assess the automatic generalisation software tools on a common basis (Burghardt et al., 2007).

## 3.6 Funded projects and research context

### 3.6.1 Basic research on automated generalisation

Since 2000 two basic research projects on automated generalisation under the co-leadership of the author are supported by the Swiss National Foundation:

- ❑ **Generalisation on Demand** (GENDEM, from 2000 until 2003)
- ❑ **Data Enrichment for GENeralisation** (DEGEN, from 2004 until 2007)

As a main result of GENDEM two PhD theses were published about "Automated Polygon Generalization in a Multi Agent System"

(Galanda, 2003) and about the "Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping" (Cecconi, 2003). Galanda (2003) investigated the automated generalisation of thematic maps. In his work the importance of structural constraints was highlighted, which had a strong impact on the proposal of the second project about data enrichment (DEGEN). Cecconi (2003) focussed on the implementation of a Multi-Scale database for on-demand web mapping. Some of the ideas of Cecconi et al. (2005) could be extended for topographic mapping during the application oriented project DRIVE (see also Section 3.6.2). For example the "linkage by id" of homologous features, was implemented in DRIVE in a modified way through the introduction of a resolution-relation class, supporting m:n relations between aggregated features as well as enabling the storage of generalisation relevant meta information such as applied operators, parameters, considered neighbourhood features, see also Bobzien et al. (2007). The mesh simplification approach presented in Chapter 8 was applied for settlement generalisation in a web mapping context by Cecconi (2003) as well.

The second basic research project DEGEN investigates the enrichment of spatial databases for automated generalisation of topographic and thematic maps. One focus was placed on the recognition of geospatial patterns and relationships between features of one resolution, also called *horizontal relations* (Steiniger et al., 2006a,b, 2007). In Chapter 5 the detection of settlement patterns will be explained utilising principal component analysis. A second category of relations proposed for data base enrichment are the so called *vertical relations*, which link features that represent the same real-world phenomenon at different map resolutions (Neun and Steiniger, 2005). They are the crucial characteristic of multi-representation databases. Vertical relations support for instance update propagation, quality and consistency checks, multi-resolution analysis functions as well as zoom functionalities for web mapping and mobile applications.

Further basic research of our group was carried out on generalisation architectures and technical frameworks enabling automated generalisation for topographic cartography, web mapping and mobile applications, e.g. Edwardes et al. (2003a), Burghardt et al. (2004a), Neun and Burghardt (2005), see also Section 3.4.

One main result will be presented in Chapter 4 (Burghardt et al., 2005c) with the utilisation of the web service concept for generalisation purposes. Especially interesting in the context of data enrichment (Project DEGEN) are the proposed category of support services for the calculation and explicit modelling of spatial data structures (Neun et al., 2006, 2008). Further categories are the generalisation operator services providing the generalisation functionality and the processing services responsible for the chaining and orchestration of operator services (Burghardt and Neun, 2006; Neun et al., pted). The generalisation service architecture is implemented both for research purposes (Edwardes et al., 2005b, 2007) as well as in a cartographic production software (Petzold et al., 2006b) as will be explained in the next subsection.

### 3.6.2 Application oriented research and development

A second category of projects carried out in our group in the past five years are industry related projects promoted by the Commission for Technology and Innovation (CTI) at Swiss Federal Office for Professional Education and Technology. Besides the transfer of scientific and research knowledge to practical applications, the industry projects helped with the identification of further research tasks. The title of the first project "**Derivation of Vectormodels**" (DRIVE) summarises the project goal of developing a generalisation platform connected with a multi-representation databases for the derivation and linkage of digital vector models at different resolutions (Burghardt and Mathur, 2004; Bobzien et al., 2005a, 2006b).

Research tasks during the project are related on the one hand to the development of core generalisation functionality as presented in Chapter 7 (Burghardt, 2005) with the smoothing of lines by snakes and in Chapter 8 (Burghardt and Cecconi, 2007) with the building typification based on mesh simplification. On the other hand investigations are carried out on the explicit modelling of horizontal, vertical and update relations in a multi-representation databases as explained in Chapter 6. Especially the vertical and update relations are used to store necessary meta information to support incremental generalisation in case of data updates (Bobzien et al., 2005b). Here a partial overlap with the research in the GENDEM project on data enrichment and relation modelling made it possible to exchange ideas on terminologies and discuss possible solutions. Furthermore some application oriented implementations are realised (Burghardt et al., 2005a). An example is the

development of a generalisation toolbox for the derivation of city maps from large scale cadastral data (Petzold et al., 2006a, 2005). With that a successful knowledge transfer from university to industry was carried out.

The second, ongoing KTI-project with title "Web based generalisation services for online maps" was set up also on conceptual ideas developed during the basic research projects. As described in the section before there are approaches and developments carried out on the topics of both web mapping and especially on the usage of web services for generalisation purposes. The acronym SerAx is made up of **S**ervices and the first two letters of the industry partner **A**xes Systems AG. The aim of the project is the creation of a generalisation platform based on a client-server architecture to provide process knowledge on automated generalisation rather than sharing geographic data only Petzold et al. (2006b). Applications and requirements exist for the production of paper maps, as well as for the dynamic creation of digital maps with application in web mapping and location-based services, see also next Subsection and Chapter 10.

### 3.6.3 New media and mobile information systems

The third category of projects comprises research on mobile information systems and adaptive cartography on small screens. With Chapter 9 (Edwardes et al., 2005a) and Chapter 10 (translated from Burghardt et al. (2005b)) two papers have been selected for this thesis, representing results from a European research and development project called WebPark (from 2002 until 2005). Within this project a platform was built to deliver location-based services in protected and recreation areas. Location-based services (LBS) denotes position dependent, personalised services, which are tailored to the individual requirements of tourists and visitors. The WebPark consortium was developed by six partners:

- ❑ two groups from industry: European Aeronautic Defence and Space Company (EADS), and Geodan Mobile Solutions;
- ❑ three academic partners: City University London, University Zurich, and Laboratório Nacional de Engenharia Civil Lissabon;
- ❑ one group from the Swiss National Park

The industry partners were mainly responsible for developing the system architecture and web mapping technology. Swiss National

Park provided content, such as animal and plant observations, route descriptions, points of interest (POI) etc. and the testing environment. The research institutes investigated questions about knowledge discovery ([Mountain, 2006](#)), the use of intelligent agents ([Mountain et al., 2003](#)) and dynamic visualisation on small displays with help of cartographic generalisation (the main topic of the Zurich group).

Research of our group was dedicated to the incorporation of generalisation functionality into the service architecture ([Burghardt et al., 2004b](#)), the generalisation of point like features such as graphical icons or animal observation data ([Edwardes et al., 2003b, 2005a,c](#)), the utilisation of hierarchical data structures such as quadrees ([Burghardt et al., 2005b](#)) as well as requirements and applications of location-based services in a broader, geographical context ([Edwardes, 2007](#)).





## 4 Generalization Services on the Web - a Classification and an Initial Prototype Implementation

Dirk Burghardt, Moritz Neun and Robert Weibel

published in  
Cartography and Geographic Information Science,  
Vol. 32, No. 4, 2005, pp. 257-268

**Abstract:** Much progress has been made in the field of web-based cartography through standards developed by the Open Geospatial Consortium (OGC). While automated access and presentation of cartographic data have been defined, the services for automated generalization are yet to be standardized. This paper aims to show advantages of applying the service concept to generalization and suggests several classification schemas of generalization services at different levels of granularity. A detailed explanation of a real implemented Generalization Service is provided. We show how software developers can make their generalization functionality available as a service and how these services can be accessed dynamically. For the implementation, the open source Java Unified Mapping Platform (JUMP) was extended to work as a framework for generalization. Generalization services could be used in different application scenarios, for instance as a middleware component extending a web map service with adaptive zooming or as stand-alone services, supporting the production of topographic maps by national mapping agencies. They may also allow the development of a common research platform, where researchers would have access to a common generalization framework.

## 4.1 Motivation

In recent years, cartography has evolved in a new direction with the application of web technologies and services. As a result of this process, maps in the web context are generally generated on demand and on-the-fly, containing more specific and tailor-made information. This contrasts with the traditional way of map making which focused on the production of static maps that were designed in advance for general-purpose usage, as is exemplified by topographic maps. Web cartography can benefit from the standards developed by the Open Geospatial Consortium (OGC) to implement web services for feature access (WFS, 2005) and web mapping services (WMS, 2005). Obviously, however, the demand for dynamic web map generation has also led to increased requirements on automated map generalization procedures. Not surprisingly, therefore, the OGC has also expressed interest into "feature generalization services" as part of their OGC Web Services (OWS) initiative (OGC and ISO, 2002). However, the specification process for such services has not had much progress so far. The map generalization research community began to develop an interest into generalization services as well, driven primarily by the desire to develop a common open research platform that would allow testing and sharing of generalization algorithms (Badard and Braun, 2003; Edwardes et al., 2003a). This has been evidenced through discussions at the meetings of the ICA Commission on Map Generalization and Multiple Representation in Paris 2003 and Leicester 2004.

There are several advantages to using generalization services in a collaborative and distributed research environment, as well as for on-demand map production. First of all, the platform independence makes the development independent from the operating system and the hardware used, which also offers application in a mobile context. Secondly, the service can be integrated in any software platform, such as web browsers and GIS or map production software. Last, the service can be accessed over the Internet or on the local desktop, provided the code of the generalization operations is made available as open source.

The objectives of the proposed paper are two-fold: 1) to present a classification of generalization services, and 2) to report on a prototype implementation of sample generalization services, as well as on experiments that were carried out to assess the feasibility of the approach.

## 4.2 Web services for spatial applications and generalization

### 4.2.1 Web services

In order to enable computers directly to access distributed data and to use services, the concept of web services has been introduced. Software-rather, computers-can read web pages, mostly in HTML, but they do not understand them. Human beings are able to read a web page and find the link or the button on which to click. To enable computers to do this without the help of a human user, web services make use of machine-processable interface descriptors and of a standardized language:

A *web service* is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Other systems interact with the web service in a manner prescribed by its description using SOAP messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards (W3C, 2004).

*Web service*

The usage of such a web service can be either fully automatic within, for instance, GIS applications or the web service could be called from the application on demand.

### 4.2.2 Spatial web services

**Human-computer interaction service** Most spatial web services are understood as data delivery services. Usually they connect to a geographic database and retrieve information from a database upon request. Examples are catalog and gazetteer services for data search or Web Mapping Services (WMS) for the visualization of spatial information (Fitzke et al., 2004). Common to this kind of services is that they are designed for end use by humans. Therefore, this category is commonly called human-computer interaction services. The approach represents the view of the Open Geospatial Consortium, which focuses on the end user. The aim is to simplify the exploitation of spatial data and the access to geo-information by the user at the end of a chain of spatial services. Often Web-based services that are connected to a database can also perform various operations on the data. Examples of such services are the

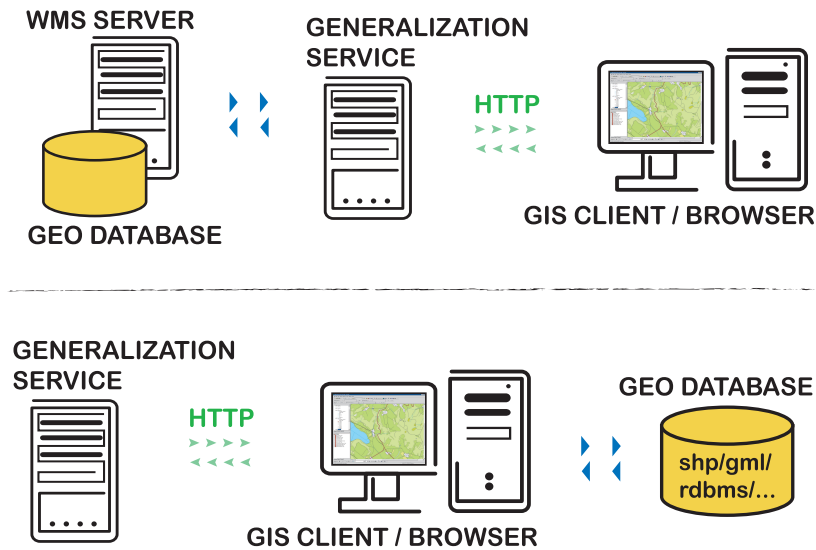
Web Feature Service (WFS), Web Coverage Service (WCS), or a route planning service.

**Computer-computer interaction service** Pure processing services, which receive, data and parameters from the requesting clients and perform operations on them, form the second main category. The possible reasons for this type of services are the availability of the algorithm only on the server due to licensing or platform incompatibilities, or the need of faster calculation power on a super-computer. This service category reflects the original purpose of web services, as given in the above definition, which includes the interoperable interaction between distributed applications, the prerequisite being a complete automation of processes.

### 4.3 Application of service concepts to generalization

In general, two concepts of generalization services seem to have a promising range of application. The potentially most widely used service would be generalization, or better, adaptive zooming in web map services. Providing a zoom function is classically the domain of Multi-Resolution Databases (MRDB). A Generalization Service could be introduced as an add-on to a web map service or as a middleware between a WMS and the client, which produces the desired resolution out of the available data. This kind of service is a computer-computer interaction service and requires fully automated solutions.

The other promising service would be the more interactive generalization services for GIS users and for map production in organizations such as national mapping agencies (NMA). In this case the service would provide functionality and calculation power to the service subscribers. It also fulfills the requirements of a common research platform (Edwardes, 2007), such that researchers have access to a common generalization framework. The advantages of such a standalone service are, for example, the ability to provide specialized or novel algorithms to the research community without everybody having to adapt their systems to the specific needs. Furthermore, the possibility exists to write specific algorithms for special computer architectures, e.g. clusters, grids or other parallel processing systems, and offering this service to the subscribers.



**Figure 4.1**  
Generalization  
services acting as a  
middleware (above)  
or as a research  
platform.

Figure 4.1 illustrates the two types of generalization services. In the case of zooming only, configuration parameters such as resolution and bounding box are sent from the client to the Generalization Service. In the case of a specialized Generalization Service, the data are sent as well. In the case of zooming, the type and structure of the data are known to the service (Lehto and Sarjakoski, 2004; Illert and Afflerbach, 2004). In the second case, this is a major problem to handle. Only standardized and valid data can be handled by the Generalization Service.

Additionally, the interactions with such services are different. While the Generalization Service runs completely automatically when it functions as a middleware, the user has (if he/she wants to have) more control in the second, research-platform, interactive scenario. These different usage scenarios will be outlined in detail later.

## 4.4 Characteristics of web-based generalization services

Generalization services take advantage of many of the general concepts underlying any web service. The outline of the conceptual and implementation characteristics, which follows, shows what a service-oriented architecture is and which techniques are avail-

able to offer generalization services to web map services or to the human users of cartographic software applications.

#### 4.4.1 Service concept and component architecture

A service is an abstract *resource* that represents a capability of performing various tasks. This abstract service has to be realized by a concrete agent. Different services can be connected with a request-based communication. This *service-oriented architecture* approach offers resources to other users in a network as independent services. Service access is achieved through a standardized approach. Loosely coupled services offer more flexibility than other system architectures. In essence, service-oriented architectures are collections of services communicating with each other. Communication can be in the form of a simple passing of data, or it could be two or more services jointly coordinating some activity. A service represents the endpoint of a connection. A service has an underlying computer system that manages the client-server connection. Service invocation is done by a *service request* to the invocable interface of the service. Upon successful operation, the service provider returns a *service response*. Errors have to be handled with *service exceptions*. These interactions are independent, and interfaces and protocols yield the underlying infrastructure. Such capabilities for service communication and interfacing are inherent in web services technology.

Generalization services can be used by higher-level services and can themselves use the functionality of other services. This is achieved through the n-tier distribution capabilities of web services. This structure implies a client/server program model. The processing of a specific application occurs over  $n$  computers across a network. These tiers usually include a data tier, business logic tier, and a presentation tier. A given computer will perform the specific tasks of a tier. Sometimes, multiple tiers can also reside on one computer, but such a configuration is usually used for testing purposes only. The remote architecture remains unchanged.  $N$ -tier applications have not only the advantages of distributing computing but, additionally, any one tier can run on an appropriate processor or operating system platform, and every tier can be updated independently.

The *component architecture* emerges from object-oriented and internet technologies. The systems components in a component-based architecture have generic interfaces through which they advertise their functionalities. This enables the dynamic loading of

the components. Components, i.e., software objects encapsulating a set of functionalities, interact with other components. Every component has an interface which conforms to a defined architecture.

#### 4.4.2 Conceptual and technical characteristics for web services in general

Every web service is a resource. Web service architectures implement a service-oriented architecture using web technologies. The following main characteristics can be defined for all web services:

- ❑ Platform independence;
- ❑ Service registry;
- ❑ Web API-interface; and
- ❑ Loosely coupled communication

The platform independence of a web service is a major advantage. This feature, also referred to as interoperability, enables a web service to be distributed over many different platforms and distinguishes it from other technologies such as CORBA or Java RMI. When browsing the web and accessing services with a browser, users expect to be able to see and use a web page on any platform or operating system. They do not care whether they are using a Windows, Macintosh or Unix/Linux computer. Web pages use a common set of standardized protocols and languages. To achieve this interoperability equally for web services, a set of already existing standards has been chosen. Web protocols ensure easy integration of heterogeneous environments. The protocol HTTP and the language XML are available for every major platform.

In order to announce the availability of a web service and to find web services, a registry is needed. The "*publish-find-bind*" principle describes this functionality. Web services are registered (*published*) and can be located through a web service *registry*. Service consumers can *find* suitable web services through a registry. These service consumers may be humans or computer applications. The registry offers a single-point access which then gives the exact service end-point of the service's implementation to the service consumer (*bind*).

The Web API is the *interface* of a service which can be called from other programs. This interface is a standards-based application-to-application programming interface which can be invoked from

nearly any type of program. In order to enable programs to bind an interface automatically, the capabilities of an interface are shown through self-description. This usually is achieved through an interface description language (i.e. WSDL). The binding of a web service specifies the protocol and data format used for transmitting messages to the invoked interface.

Web services are *loosely coupled*. The systems pass XML messages, usually over the HTTP protocol, to each other via their Web API. The Web API interface is the abstraction layer which yields the real communication and makes the connections stable and flexible. The HTTP protocol and the XML language can transport nearly any type of data. For instance, binary arrays can be converted into ASCII and then packed into an XML document. This approach makes the Web Service concept very open and usable for many different purposes.

#### 4.4.3 Advanced characteristics for generalization services

Depending on the usage concept-either as a stand-alone service (e.g., research platform) or as middleware-different requirements of service registry and service invocation exist. The middleware concept often needs no registry because mostly it forms a combined system together with the data source (e.g., WMS). So, in this case the middleware would form some kind of service endpoint for the data source, and the Generalization Service is bundled with data access. The success of generalization services as a common research platform or stand-alone Generalization Service, however, is very much dependent on the existence of a service registry (see below). The service registry has to be the single point access to all available services. These "Yellow Pages" for generalization services must know where the interface description and the service endpoints of the desired service can be found. The concrete architecture of such a registry will be described later in this paper.

Like the service registry, *service invocation* also depends on the usage concept. A middleware concept of access to generalized data would only need to transfer parameters to the service, while a research platform or a stand-alone Generalization Service also needs functionality to upload data. All these concepts also can have different ways of interaction; the classical human-computer interaction is based on forms (in web pages) or application plug-ins. This



interaction can be fully transparent to the user or be hidden in a cartographic system without any user control or interaction.

The interface of, and the encoding of geo-features by, a Generalization Service (which supports the uploading and treatment of individual users' data) both have special requirements with respect to their geometry and semantic. In a form-based web page environment with full human interaction, transfer formats such as Shapefiles or GML would be suitable. But they have to respect the format needed by the service. When using a plug-in in a cartographic application, the communication logic of the plug-in has to translate the application's internal concept of geo-features to a common format which can be understood by the service.

Generalization of more than one feature and complex generalization services need two additional features which are not built into web services. These features are a capability to maintain the program's *execution state* and support for *atomic transactions*. Maintaining the state of the execution of an algorithm is very important when client interaction with the service at run time is desired. Once the client has uploaded the initial parameters (and possibly data), the service starts the execution of the algorithm. If the algorithm needs additional data, parameters, or user input, a request is sent back to the client. While this communication takes place, the service has to maintain the already entered information, data, and the already calculated changes. As web services are based on loosely coupled state-less communication, this maintenance requirement is accomplished by such technologies as session management via session IDs or cookies which are built into many web servers. The notion of transactions is fully compatible with the transaction concept in database systems. While executing a complex algorithm on multiple features in one data layer, the data set is only changed if no error occurs. Otherwise, a roll-back is done and no changes are committed. This feature is important to maintain the original data set's consistency in generalization.

The state maintenance and the concept of transactions become extremely important when advanced *service control* and, especially, the *chaining* of several services is desired. The chaining of different services with their algorithms is one important step towards a "generalizeMap()" operation which generalizes an entire map and delivers a more-or-less ready-to-use product. In middleware systems, where the objective is fast display of generalized map objects, the steps of the processing chain are opaque. Here the quality is less important than speed and stability. In map production and

research service chaining, interaction with the user is a must in order to obtain high quality results. The usage concept is implemented using a plug-in, because a form-based solution (e.g., in a browser) does not offer the flexibility and the support for long-time connections.

## 4.5 Generalization services categories

### 4.5.1 Categorization based on OGC/ISO architecture model (02-025)

As discussed above, the involvement of humans in spatial services has led to their distinction between human-computer and computer-computer interaction services. The OGC/ISO architecture model refines these two main service categories with regard to the following:

- ❑ Processing services;
- ❑ Model/information management services;
- ❑ Workflow/task management services;
- ❑ Human interaction services;
- ❑ Communication services; and
- ❑ System management services.

Services belonging to these different categories can be applied usefully to the generalization process, as we will now show. *Processing services* encompass services that perform large-scale computations, which are needed when a generalization process is fully automated, for instance, on a complete map image or on generalization sub-tasks which can be completely separated from each other, such as text placement. Typically, processing services do not include capabilities for providing persistent storage of data. With regard to our application scenarios, processing services were needed to implement on-the-fly generalization, such as adaptive zooming. Between data access through a Web Feature Service and map presentation with the help of a Web Mapping Service, the Generalization Service runs completely automatically as a middleware, adapting information to the screen size of an output device, used symbolization, and map content. In this context, performance is very important, so lower quality of the Generaliza-

tion Service will be accepted, i.e., simple selection and simplification operations are carried out in real time, omitting more time-consuming, context-dependent generalization operations. Lower quality maps usually are sufficient for display on mobile devices, where the image-loading time affects the user's look-and-feel capabilities and where limited screen resolution limits the display of detail.

Other application scenarios use the Generalization Service as support for interactive map production and user-controlled semi-automated derivation of multiple representations at smaller scales. The *model and information management services* category from the OGC/ISO architecture model can be seen as a collection of these kinds of services, which allows the development, manipulation, and storage of metadata, conceptual schemas, and datasets. In this type of application the main control is with the user. *Workflow services* help to define, invoke, and control the status of service chaining. *Human interaction services* allow interaction during the generalization process, for instance, while deciding which object classes have to be generalized or selected and to parameterize Generalization Service operators. The OGC/ISO architecture model suggests two additional categories, one for *communication services* to encode and transfer data across networks and one for *system management services* which include, for instance, management of user-access privileges. All these kinds of services can be advantageously combined in a generalization research platform.

#### 4.5.2 Default and advanced generalization service categories

Another way of categorizing generalization services is to apply the OpenGIS Geography Markup Language (GML) Implementation Specification, which differentiates between GML core schema elements and GML application schema elements. Default generalization services as one category deal with GML core schema elements and could be used with any GML dataset. In many cases, for example when simplifying single area objects (e.g., the geometry of a building) or line objects (e.g., the geometry of a river or street), application requirements may be simple, and default generalization behaviors may suffice to meet those requirements. Advanced generalization services encompass services supporting GML application schema elements. These elements enable the modeling of objects by means of attributes and object compositions.

Default generalization services contain generalization functionality which operates at the micro level; i.e., generalization operations are carried out on single objects, while context dependency is not be considered. According to Ruas (2000), micro objects generalize themselves or react to orders for contextual operations given by (superordinate) meso objects. Using the typology of McMaster and Shea (1992), the following operators can be applied on single objects: simplification, smoothing, geometry type change, collapse, enhancement, and selection based on geometry. Default generalization services enable resolution reduction by means of a single geometrical operation (e.g., line simplification) and simplified modeling (e.g., centerline representation of roads and streets).

Advanced generalization services consider attributes, symbolization, and spatial context through neighboring objects. Hence, their generalization operations have to deal with groups of objects, so called meso objects. Meso objects (or meso agents) generalize themselves when they perform contextual operations (Ruas, 2000). Advanced generalization services implement such contextual generalization operators of typology (McMaster and Shea, 1992) as context-dependent selection, aggregation, typification, exaggeration, and displacement of map objects, by taking into account the geometry, semantics, and attributes of the objects involved.

### 4.5.3 Hierarchical categorization

Extending the idea of default and advanced generalization services, a hierarchical breakdown can be made which distinguishes between the following categories:

- ❑ Generalization support services (e.g., services for buffering or for creating a topological data structure, a skeleton, or a constrained Delaunay triangulation);
- ❑ Generalization operator services (e.g., services for line simplification, line displacement, area aggregation); and
- ❑ Generalization process services (e.g., services for automated orchestration, services for evaluation of generalization results).

The first category of services is intended to support the generalization process and the generalization operators and thus represents the lowest level of services. Examples are services for creating a topological data structure or services for creating a constrained Delaunay triangulation. The result of such a service is

additional cartographic information which can be optionally stored also in a Multi-Resolution Database (MRDB). Support services enrich the data being generalized; they make information explicit, by representing common structural properties such as neighborhoods or proximity relations and alignments which can be usefully exploited by generalization operations (Neun et al., 2004).

The second service category delivers the functionality of standalone generalization operators. Several typologies of such generalization operators are suggested, for instance by McMaster and Shea (1992). Examples are services for simplification, smoothing, aggregation, amalgamation, merging, collapse, refinement, exaggeration, enhancement and displacement. These generalization operator services can be further subdivided for point, line, and area objects and specialized depending on object classes. It is obvious that rivers, borders, and railway lines have to be generalized in a different way, despite the fact that all are line objects. The generalization operators are offered in an interactive mode, with the user selecting appropriate generalization operators/algorithms, as well as setting the control parameters of the algorithms.

The third level of services, generalization process services, use services from lower categories and enable the control and orchestration of generalization operators. Examples are the meso agents described above in connection with advanced generalization services (Ruas, 2000). Automated control of the generalization process presently receives ample attention as a research topic. Besides agent-based modeling, combinatorial and continuous optimization approaches have been proposed in the literature. Simulated annealing (Ware et al., 2003), which is a combinatorial optimization approach, allows the selection of generalization operations which are controlled by assigning costs to each operation. Continuous optimization approaches include the finite element method (Højholt, 2000), snakes or elastic beams (Burghardt and Meier, 1997; Bader, 2001; Galanda and Weibel, 2003), and least-squares adjustment (Harrie, 1999; Sester, 2000).

The latter methods are primarily used to control generalization operations of continuous nature, such as displacement or smoothing. All approaches mentioned so far, however, are still quite a way from a perfect modeling of the overall map generalization process. An interim mechanism for controlling the generalization process is however available for service chaining in the OpenGIS Web Services Architecture (OGC and ISO, 2002), which

can be used depending on purpose and map complexity. This mechanism consists of:

- ❑ User-defined (transparent) chaining: The human user manages the workflow for complex generalization processes for which no adequate process modeling exists yet.
- ❑ Workflow-managed (translucent) chaining: The human user invokes a workflow management service that controls the chain, and the user is aware of the individual services. For medium-complexity generalization processes this can be specified as workflows.
- ❑ Aggregate service (opaque): The human user invokes a service that carries out the chain, with the user having no awareness of the individual services. This approach is suitable for relatively simple, sequential generalization processes.

## 4.6 Implementation

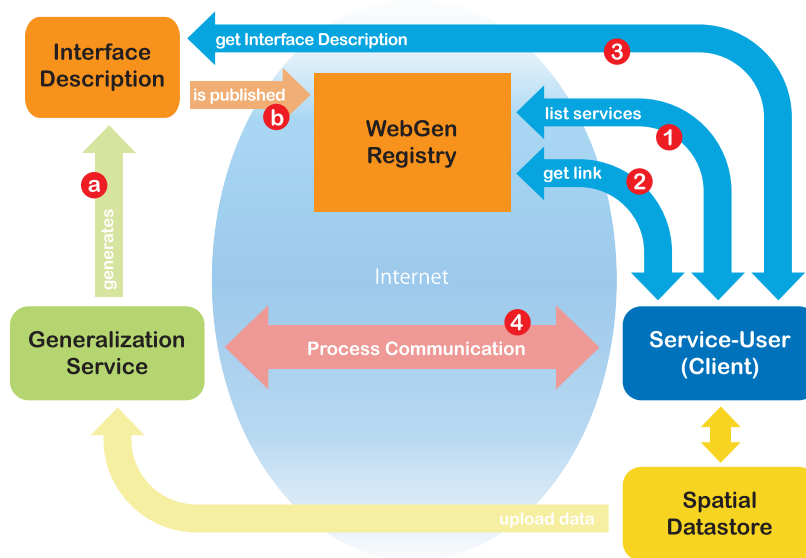
The prototype implementations of the generalization services reported here include examples of rather simple services (e.g., Douglas-Peucker line simplification, building simplification) where spatial context is not considered. The objective is to show the feasibility of the service-based approach and to describe the minimum set of components needed to run the generalization services over the Internet. To implement the generalization services we use an open source framework for mapping - JUMP Unified Mapping Platform (<http://www.jump-project.org>) - which delivers standard functionality for reading and writing files (Shape, GML) as well as modifying and visualizing cartographic data. JUMP is written in Java, which enables easy application service provision over the Internet. The usage of the framework's functions has been enhanced by including the JUMP libraries in Java servlets running in such servlet engines as TOMCAT. These servlets may contain algorithms with the generalization logic or controls for external generalization libraries.

### 4.6.1 Generalization services registry

The central point for accessing and publishing generalization services is the registry. Similarly as with the yellow pages, the registry is used to locate available generalization functionality. It

would make sense for a large independent organization such as the ICA Commission on Map Generalization and Multiple Representation to host such a registry database.

Providing a generalization service to, for instance, a research group, involves several steps (see Figure 4.2). The first task is to create an interface description containing the parameters of the generalization service and the service endpoint; the description will include a URL address where the service can be accessed. The second is to create the registry database. Once the interface description is published in the registry database the community can access the service.



**Figure 4.2**  
A registry for  
generalization  
services.

Clients looking for generalization services can use the registry to find what services are available. After selecting the desired generalization service, the client obtains the link to the interface description of the service. When accessing the interface description the user knows the endpoint of the service, as well as the number, names, and types of the parameters. With this information the client plug-in can automatically create a user interface for the selected service, allowing the specification of generalization parameters. Having a central access point is one advantage of a registry. Additionally, the registry automatically identifies service endpoints, so there is no active change needed by the service users if the service provider changes the location of the software on the

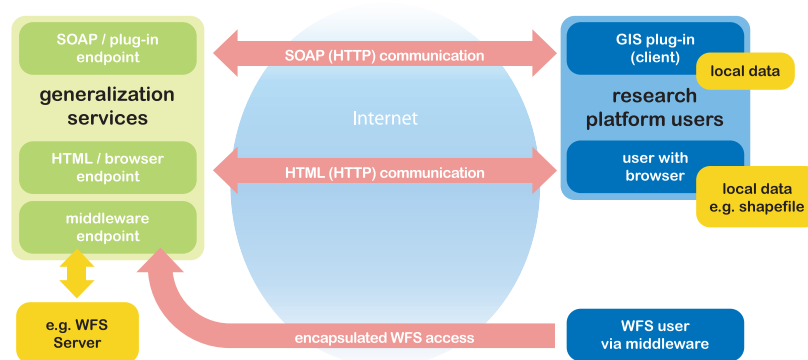


server. Finally, process communication is responsible for the transfer of parameters and data between user and service provider.

#### 4.6.2 Process communication and feature metadata (encoding and upload)

The middleware and research platform concepts require different ways of communication for uploading (data input) and downloading (output). Figure 4.3 illustrates three different ways of implementing process communication. The research platform concept may be implemented as a form-based web page in a browser or as a mapping software plug-in. The browser upload offers only an HTTP file upload. The selected file is encoded automatically by the browser and has to be decoded on the server. The browser upload option offers the possibility to upload and process shapefiles, i.e., the user downloads a newly created shapefile as an output in the browser, together with the generalization algorithm.

**Figure 4.3**  
Three ways to  
implement process  
communication.



The implementation as a plug-in to the client GIS offers the possibility of encoding data directly out of the application. As a result one can choose a better suited format for the transfer. In the plug-in we tested, the geo-features are encoded in a GML compatible format directly into a SOAP message (Simple Object Access Protocol), then transferred to the server. Other formats (e.g., a binary format) could be implemented, possibly with SOAP envelopes and, for instance, MIME encoding. The output of the server is in the same format as the request and can again be decoded to the client and displayed there.

The concept of a middleware layer between the data source and the user (mostly in conjunction with a browser) introduces a third



possibility for getting the data to the generalization service. In this case the client would simply send the URL of the data source (e.g., a WFS) to the generalization service as a parameter in the call. The server itself would then access the data source, process the data, and send it back to the user (Sester et al., 2004).

The metadata specifications of the geo-features are treated in the following way. Every feature can have an arbitrary number of attributes. The number, name, and type (spatial or non-spatial) of the attributes needed by the algorithm are specified in the interface description. Generalization data are always geometrically dominated; this means that every feature has at least one geometry plus possibly other attributes. The interface of a road generalization algorithm may require a geometry for each road and the road class (highway, major road). Listing 4.1 shows an example of such a schema specification. The users of the service select the columns with the desired geometry and the road class in their local data set, on their client computer. To transfer the data, the client then automatically assigns the name from the schema to the selected column. In this way, the server identifies the attributes it can use. Other attributes in a local dataset will simply be ignored, not deleted.

```
<schema>
  <attribute>
    <name>geometry</name>
    <type>GEOMETRY</type>
  </attribute>
  <attribute>
    <name>class</name>
    <type>STRING</type>
  </attribute>
</schema>
```

**Listing 4.1**  
*An example of a  
feature schema.*

All the encoder and decoder methods necessary for the browser upload and plug-in have been implemented in Java. They form the communication framework which can be used for client-server communication.

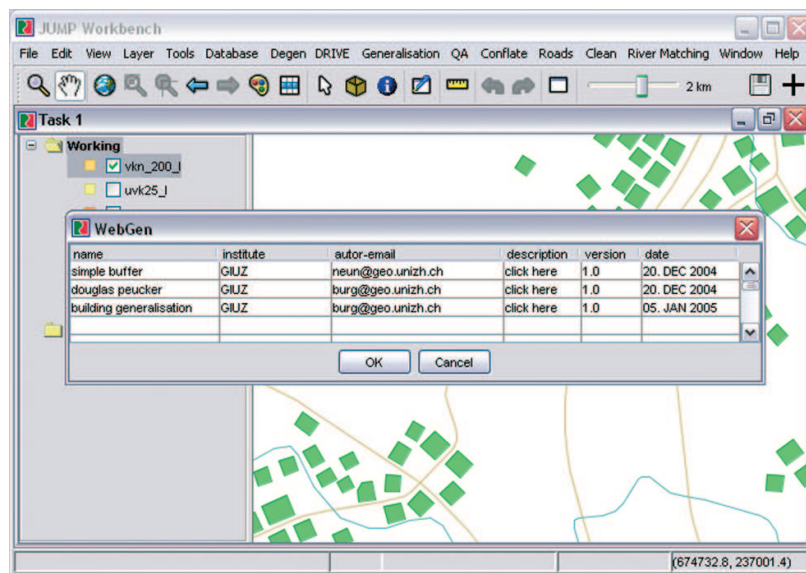
### 4.6.3 Service invocation and client implementation

On the client side, both a browser-based generalization service and a plug-in for the JUMP Unified Mapping Platform have been implemented. Other plug-ins for platforms such as ArcView are planned. The client implements an easy to use GUI (Graphical User Interface) for the service user. The browser example works with standard HTML pages on every major browser platform. The

user gains access to such a service via a start page which includes user authentication with a password to activate access. The start page, which lists all available services, is dynamically created with information from the service registry. After selecting a particular generalization service, a new, dynamically created page is presented to the user, which allows him/her to enter the parameters for the algorithm and upload a shapefile from his/her local system.

The JUMP plug-in does mostly what the browser solution does, but it integrates seamlessly into the software so that the user does not have to quit the application, and even does not necessarily notice a big difference between using a local or remote algorithm. The plug-in integrates into the JUMP menu bar. The first time after the plug-in has been installed the user will be prompted to enter the URL of the registry. With this URL, the plug-in automatically looks for all available generalization services every time it is started. The result of this query is displayed to the user in the form of a selection list (see Figure 4.4).

**Figure 4.4**  
Available services  
(from the  
Generalization Service  
registry).



From this list of available services, the user selects the desired Generalization Service. The user then receives an entry form for the corresponding algorithm's parameters. The configuration of all those entry forms is dynamically created from the interface description. An example of such a simple interface description for the Building Simplification algorithm is shown in Listing 4.2. The interface description is in XML. The important parts of the interface description (Listing 4.2) are highlighted.

```

<?xml version=1.0 encoding=UTF-8?>
<webgen>
  <name>building generalization</name>
  <method>buildingGenNew</method>
  <endpoint>http://www.geo.unizh.ch:8080/neun/servlet/
  GenHandlerXML</endpoint>
  <description>This algorithm simplifies all buildings in
  a layer and returns a layer containing the resulting
  buildings!</description>
  <config>
    <layer>
      <schema>
        <attribute>
          <name>GEOMETRY</name>
          <type>GEOMETRY</type>
        </attribute>
      </schema>
    </layer>
    <param>
      <name>min edge length</name>
      <type>DOUBLE</type>
      <description>Minimum Edge Length!</description>
    </param>
  </config>
</webgen>

```

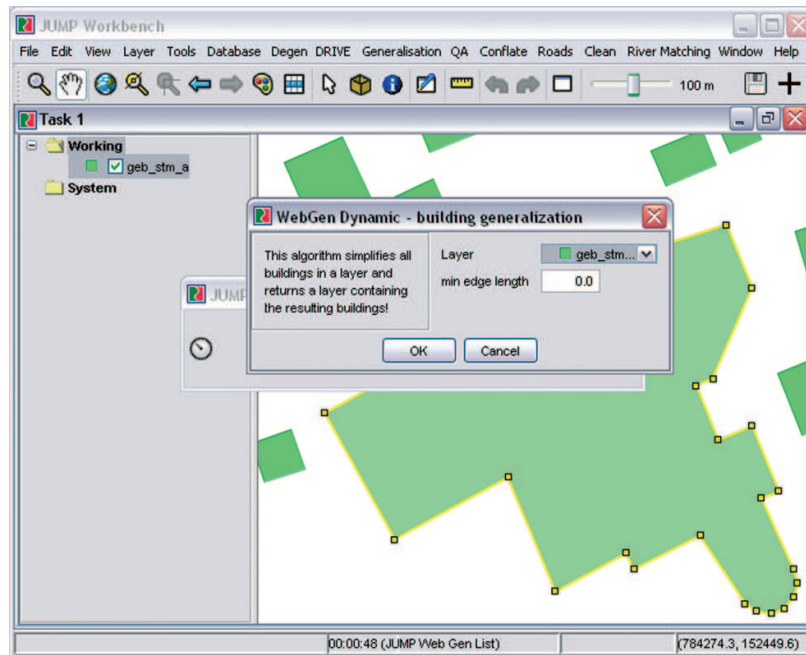
**Listing 4.2**  
The interface  
description in XML.

The entire description is in one XML container. The tag `<method>` specifies the generalization algorithm which has to be used on the server and passed to the server proxy. The tag `<endpoint>` specifies the URL of the server proxy. The tag `<layer>` specifies the minimum schema (metadata) the algorithm needs. In the example (Building Simplification algorithm), only geometries (`<name>` and `<type>`) are needed. Other attributes can also be contained in the layer but they will be ignored. For another algorithm, such as the road re-classification algorithm, a second attribute indicating the class of the road would be needed. The other parameters which are needed by the algorithm are indicated by the `<param>` tag. There can be as many parameters as needed. In the example, only one parameter—the minimal edge length (type DOUBLE)—is needed.

Figure 4.5 shows the automatically generated entry form for the Building Simplification algorithm. After selecting the layer with the geometries and entering the tolerance, the data are transferred to the server and processed. The resulting geometries are sent back to the client and displayed in a new layer in JUMP.

The implementation example assumes a near real-time execution of the generalization service. The client is kept in a blocked wait mode, and the generalized data are delivered back to the client as soon as they become available. Execution time is limited by client and server timeout settings and users' patience. For more time-consuming operations, a system which informs the user

**Figure 4.5**  
*Algorithm interface  
 generated  
 dynamically from an  
 interface description.*



when the service is finished would be preferable, and it is possible. The advantage of such a system is that clients would not be blocked during the processing of their requests, allowing them to perform other tasks. Such a system has not been implemented so far, however. For the middleware concept, only the real-time execution is of interest because users usually want to see the map very quickly. The (simple) algorithms implemented so far have all performed very fast, so the bottleneck was not the computing time but the network transfer time. In the middleware concept, assuming a fast connection between data source and the generalization service, this is negligible.

## 4.7 Conclusion

In an attempt to stimulate discussion about generalization services in the generalization research community, this paper focused on the minimal requirements for successful client- and server-side implementation of generalization services. It was shown how developers and researchers can make their generalization functionality available by means of an interface description, and how possible users of these services can find them through a registry database.

As an example, a simple generalization service for line and building simplification was implemented and accessed dynamically from the open source Java Unified Mapping Platform. To show the advantages of generalization services, two different application scenarios were proposed. The first scenario implements the processing service as a middleware solution which enables on-the-fly generalization for tasks such as adaptive zooming. The second application scenario focuses on interactive generalization in support of interactive map production and user-controlled semi-automated derivation of multiple representations at smaller scales.

The limitations of our solution are that no symbolization was considered, and only context-independent generalization services were implemented. Also, no user or session management was developed. This would be needed for longer computations. Further research has to investigate ways of interactively chaining generalization services, which is related to the yet largely unsolved problem of automated orchestration of generalization operators. To take full advantage of distributed services and grid computation, ways have to be found for separating and distributing generalization tasks. The partitioning of a generalization task based on an identification of independent problems (e.g., the generalization of building blocks surrounded by streets) or the subdivision of a map area with solving boundary problems are within our reach.



## 5 Usage of Principal Component Analysis in the Process of Automated Generalisation

Dirk Burghardt and Stefan Steiniger

published in  
Proceedings of 22<sup>nd</sup> International Cartographic Conference,  
A Coruña, Spain, 2005, CD-ROM

**Abstract.** Current generalisation methods use cartographic constraints for the analysis and evaluation of map situations, as well as for the selection of generalisation operators. Despite the importance of constraints for the whole generalisation process, little research has been carried out about the relationships among constraints and their interdependencies. The aim of this paper is the investigation of such relationships with the help of Principal Component Analysis (PCA) and a model called constraint space. Three applications are presented based on the usage of PCA for the generalisation process. First, the homogeneity evaluation of building alignments, second the detection of settlement types from building datasets, and third the identification of special buildings akin to outlier testing.

## 5.1 Introduction

The process of map generalisation has been modelled in the past by several approaches, initially, through applying simple batch processing, improving this by condition-action modelling, leading to the currently favoured constraint based approaches. The underlying concept changed from modelling of the generalisation process through static workflow descriptions, to a model where only the start and endpoints are described by constraints, while the solution path in between is flexible. Constraints allow an evaluation whether the map is cartographically satisfactory or at least whether the cartographic presentation improves during the generalisation process.

The difficulty of this approach is the orchestration of generalisation operations by prioritisation and weighting among different and potentially conflicting constraints. This affects the selection of sequences of generalisation operators (often called plans), which is typically realised manually and is therefore often rather arbitrary or subjective. To formalise relations between constraints we introduce a model called constraint space, which is derived from standardised constraints. This model allows the investigation of dependencies between constraints and a reduction to the most important components. The method used here is principal component analysis (PCA).

Objects or groups of objects are placed inside the constraint space depending on their cartographic properties. The distance of the objects from the origin is a measure of the severity of cartographic conflicts. Objects at the origin are not violating any cartographic constraints. The constraint space allows us to identify similar cartographic situations for objects or groups of objects, which can then be generalised with the same generalisation operators and parameters. An evaluation after generalisation can be applied to generate a probability of successfully used generalisation operators. With this technique the system will be capable of learning how to generalise, depending on the position of the objects inside the constraint space.

## 5.2 Constraint space and Principal Component Analysis

### 5.2.1 Constraints and measures

The concept of cartographic constraints has been introduced from computer science to map generalisation by [Beard \(1991\)](#). Constraints received special importance in cartography through the



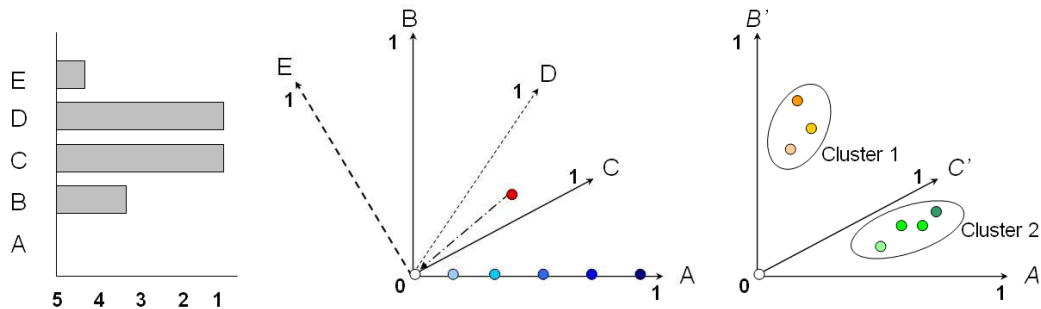
application of intelligent agents in the area of automated generalisation (Ruas, 1999). In comparison to rules, constraints are more flexible, because they are not bound to a particular generalisation action. Following the results from the AGENT project (Barrault et al., 2001) constraints designate a final product specification on a certain property of an object that should be respected by an appropriate generalisation. In the agent-based generalisation system created in AGENT, constraints are described as a collection of values and methods like goal value, measures, evaluation method, list of plans, an importance and a priority value. While measures only characterise objects or situations, without considering cartographic objectives, constraints evaluate situations with respect to the formalised cartographic objectives. Thus, the constraints check whether the objects or situations are also in a cartographically satisfying state. In this sense measures are a subset of constraints. Despite the fact that several taxonomies of generalisation constraints have been suggested (Weibel, 1996; Harrie, 1999; Ruas, 1999), the selection and also the weighting of constraints is rather subjective and arbitrary. Our model of the constraint space, which will be introduced in the next paragraph, delivers a tool to make these selections on a formalised basis.

Automated generalisation can be seen as an iterative process between conflict analysis and conflict resolution. Both, analysis and resolution, are intimately connected with constraints since the identification of conflicts and the selection of generalisation operators to remedy conflicts is based on constraints. The difficulty comes from the fact that cartographic situations are typically connected to a set of constraints that are partially in conflict with each other. Examples are the constraint "preserve minimal distances between objects", which works against the constraint "maintain positional accuracy" or the need to "reduce details" versus the constraint "maintain the original shape as much as possible". The goal of automated generalisation is to find a good compromise between all these conflicting constraints. Before the generalisation can be carried out the constraints have to be prioritised. The model of a constraint space presented below can support these distinctions, because it allows the investigation of relationships between constraints.

### 5.2.2 Constraint space

The model of the constraint space is based on the generalisation state diagrams (5.1, left), which were previously used in the context of agent modelling for generalisation (Ruas, 1999; Barrault et al., 2001). The axes of this  $n$ -dimensional space represent  $n$  cartographic constraints with their degree of satisfaction. The axes

are scaled to the interval  $[0, 1]$ , whereby a value greater zero means that the constraint is violated. The constraint violation values are equivalent to the so-called severity in the agent model of [Barrault et al. \(2001\)](#). In Figure 5.1 (middle) the brightness of the dots becomes darker depending on the distance from the origin, which should illustrate that the distance is a measure of conflicts. When working with a constraint space, a distinction has to be made between the creation and analysis of the constraint space on the one hand and the placement and classification of cartographic objects inside the constraint space on the other.



**Figure 5.1**  
Evaluation of states  
with constraint  
severity (left) and  
equivalent  
presentation with  
standardised (middle)  
and reduced  
constraint space  
(right).

The first task is the creation of a suitable constraint space for a given cartographic situation. The cartographic situation is defined by the spatial extent and the involved object classes. Dependent on the situation the cartographic constraints are selected, e.g. thematic maps have to satisfy different constraints than topographic maps, and buildings should follow partially other constraints than streets. The derived constraint space can be investigated with representative test data to detect correlations between constraints. The constraint space can be simplified by reducing the number of dimensions, if the correlation of constraints is based on a description of the same cartographic phenomenon. Hence, for instance, similar shape constraints can be replaced by one. The technique that can be used for this kind of operation is called Principal Component Analysis and will be introduced in the next section.

After the creation of this constraint space every cartographic object has its dedicated place in this space, depending on the satisfaction of its constraints. A classification can be carried out to identify objects with similar cartographic constraint values, indicated in Figure 5.1 (right). The idea is to generalise objects with the same operators if they are situated next to each other in constraint space. In Section 5.4.1 on settlement characterisation sev-

eral ways of classification are described in detail. It is also possible to apply constraint spaces for the characterisation of groups of objects, such as alignments of buildings or islands. Here constraint spaces of the individual objects can be interpreted as sub spaces. A comparison of the group positions inside the constraint space before and after application of several generalisation operators allows the derivation of probabilities for a successful usage of certain generalisation operators on certain cartographic situations. Step by step the system can learn which sequence of generalisation operators may be applied with respect to a high probability of success obtained from the position inside the constraint space.

The model of the constraint space has several applications:

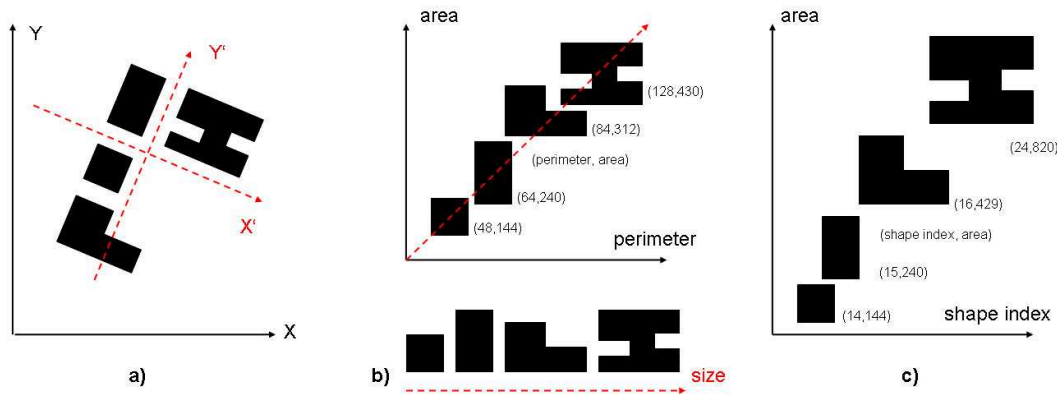
1. Correlation analysis of constraints based on a representative test dataset
2. Reduction of variables or constraints
3. Detection of outliers by position evaluation in the constraint space (an example of the identification of special buildings is given below)
4. Grouping in constraint space by characterisation of cartographic objects (e.g. settlement type detection)

The last point refers to another property of the constraint space, which has to be investigated further. The constraint space as well as the constraints can be scale dependent. Thus, some constraints are constant only for a given scale range. This has implications also on the correlations of the constraints. The next section introduces the multivariate method of Principal Component Analysis to identify correlations between constraints.

### 5.2.3 Principal Component Analysis

The central idea of Principal Component Analysis (PCA) applied to cartographic generalisation is the identification of correlative relationships between constraints or measures and the reduction of measures and constraints onto the essential components for a given data set. Every cartographic object is characterised by its position inside the  $n$ -dimensional constraint space, which could be described by an  $n$ -dimensional vector. The mathematical background of PCA is the so called "Principal Axis Transformation", which helps to identify a better suited basis for a set of given points or vectors. First, the origin of the coordinate system is moved to the centre of gravity of the points and second the coordinate system is rotated in the direction of highest variance of the point data.

Subsequently the second main axis is rotated orthogonal in the direction of the remaining main variance. This process is repeated until a new, possibly lower dimensional basis is created (see Fig. 5.2).



**Figure 5.2**

Left - "Principal Axis Transformation" with simple rotation to create a better adapted coordinate system; middle - reduction of two correlated components (*area*, *perimeter*) into one principal component (*size*); right - correlation between *area* and *shape index* ( $\frac{\text{perimeter}}{(2\sqrt{\pi \cdot \text{area}})}$ ), due to the fact that smaller buildings are generally simpler than larger ones.

The result is a better suited coordinate system such that the directions of the coordinate axes correspond with the highest variances. The first principal component is the combination of variables that explains the greatest amount of variation. The second principal component defines the next largest amount of variation and is independent (orthogonal) to the first principal component. There can be as many principal components as there are variables. The main use of PCA is to reduce the dimensionality of a vector space while retaining as much information (a high degree of variance) as possible. It computes a compact and optimal description of the data set.

Figure 5.2a) shows an example of Principal Axis Transformation which has been applied to create a better suited coordinate system. The orientation of the buildings with reference to the transformed coordinate system is axis-parallel. In Figure 5.2b) a strong correlation between building perimeter and building area can be seen, which has been expected for normal buildings and was also obtained from the PCA. Hence, the two measures *area* and *perimeter* can be replaced by one, the synthetical *size* variable. The right hand figure shows another correlation between size and shape. The reason might be that smaller buildings are in reality usually simpler than larger ones. Another reason may be the fact that smaller buildings must be more simplified than larger buildings at a scale around 1 : 25'000, because the minimal size constraint suppresses very small edges. Therefore the

small buildings are presented as squares, while the larger ones can be visualised with more detail. In this case a reduction of the two components is not recommendable, since this correlation is not constant across scale changes (see also Section 5.3.2).

## 5.3 Homogeneity of groups - subspace of constraint space

### 5.3.1 Constraints to evaluate homogeneity of building alignments

Before we can start the investigation of the constraint space, which characterises the homogeneity of building alignments, we have to create object groups in an automated way. Our approach starts with the restrictive assumption that most building alignments are situated in the neighbourhood of linear objects such as streets, roads or rivers. Based on topology and geometry the reference lines for possible alignments can be derived from line segments between crossings. Alignment candidates are selected from buildings within a given distance from the reference lines. Then, the base line of the alignment can be derived from the reference line (e.g. street) through parallel translations whereby the sum of distances between the centre of gravity of alignment candidates and the base line are minimal. After this preselection of alignment candidate buildings the quality of these alignments has to be evaluated. Several criteria were proposed in the literature (Christophe and Ruas, 2002; Ruas and Holzapfel, 2003; Li et al., 2004). We have chosen 15 measures, listed in Table 5.1, from four categories - *size* (3 measures), *shape* (5), *orientation* (3) and *groupcharacteristics* (4).

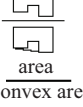
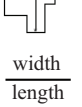
These measures are used to calculate homogeneity constraints of building alignments. The homogeneity constraints are defined by minimal variation of the measure values between buildings of one alignment. Groups are more homogeneous if the values are similar for all buildings belonging to the alignment. The process of calculating the group homogeneity consists of the following steps:

1. Calculation of measure values for every building (group measures are based on alignments).
2. Determination of variations for every measure value related to one group; e.g. calculate a maximal deviation from mean values  $\sigma_{max}(A) = \max(A_{max} - A_{mean}, A_{mean} - A_{max})$ .
3. Comparison of the variation between several groups.

For a better comparison of the variation between groups a percentage value is calculated from the deviations, which is additionally

M1 - Size	Area of building
	Perimeter of building
	Length of diameter of building
M2 - Shape	Shape index
	Building concavity
	Compactness
	Building elongation
	Fractal dimension
M3 - Orientation	Angle between longest axis of minimum bounding rectangle and horizontal axis
	Angle between longest axis of minimum bounding rectangle and base line
	Smallest angle between two axes of minimum bounding rectangle and base line
M4 - Group characteristics	Distance between centre of gravity and base line
	Nearest distance from building to reference line
	Distance between centres of gravity of neighbouring buildings
	Shortest distance between neighbouring buildings

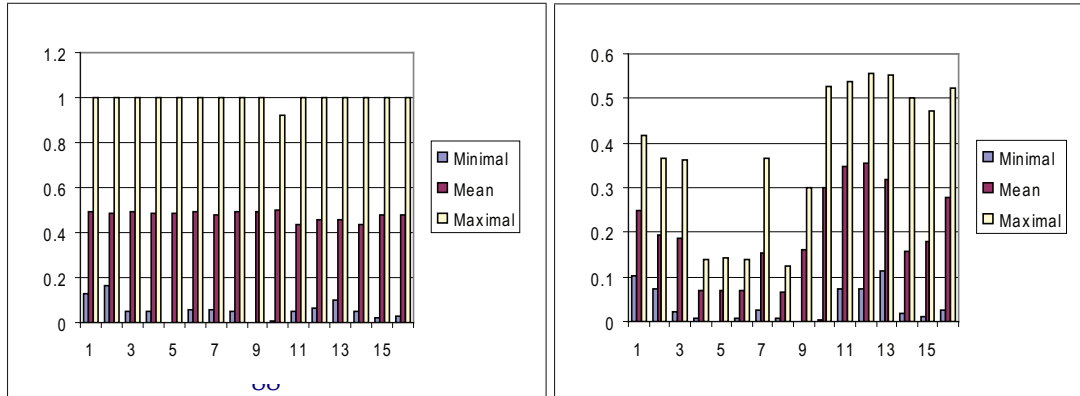
Shape index  
 $m_s = p / 2\sqrt{\pi \cdot a}$   
 Compactness  
 $m_c = 1 / m_c$

Building concavity 	Building elongation 
Fractal dimension $m_F = 2 \cdot \ln p / \ln a$	

**Table 5.1**

Measures to characterise building alignment candidates.

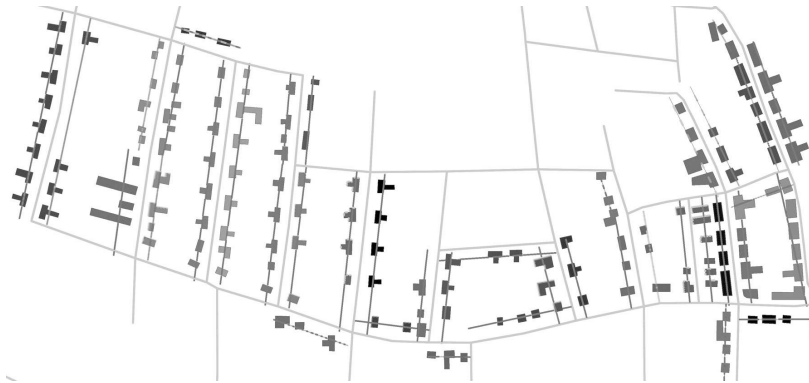
standardised to an  $[0, 1]$ -interval. The percentage values are related to mean values for every group (e.g.  $d_{Amax}/A_{mean}$ ) or fixed values (e.g.  $dO/\pi$  for orientation). The standardisation is different for the various measures since either the percentage values of deviations can be higher than 100% (e.g. for *size* deviation) or not (e.g. for *orientation* deviation). An alternative suggested by Boffet and Rocca Serra (2001) is to divide the standard deviation for every measure and group by the maximum value of the measure for every group. The differences can be seen in Figure 5.3.

**Figure 5.3**

Mean, minimal and maximal deviation values of 16 measures for 40 alignments.

Finally for every group the average of the standardised deviation of all measures is calculated and visualised. In Figure 5.4 less homogeneous building alignments are shown in lighter grey. The

information on maximum deviation of measure values for every group and individual buildings could be used to improve the quality of the alignments. Buildings which have maximum deviations could be excluded and the homogeneity value re-calculated anew for the group.



**Figure 5.4**  
Visualised  
homogeneity of  
building alignments.  
Inhomogeneous  
alignments have  
lighter grey values.

### 5.3.2 Dependencies among measures

The measures selected from the literature gave acceptable results for the evaluation of the homogeneity of building alignments. Nevertheless we were interested to find out which shape measures are similar and whether they all are indeed needed. Principal Component Analysis can be applied to investigate the chosen measures. One result of the PCA are correlation values between all measures. A further result is an ordered number of principal components, which allows the reduction of the original measures. The order of principal components is based on combinations of variables that explain the greatest amount of variation. Table 5.2 shows the correlation values between all pairs of measures.

The matrix is symmetric and rows and columns contain the measures of Table 5.1 in the same order. As we can see there is a strong correlation between all measures of *size* category (M1: *area*, *perimeter* and *length diameter*), because the correlation values are higher than 0.80 (grey cells in the upper left part). The correlation values within the other categories vary (grey cells). In the *shape* category (M2) exists a 100% anticorrelation between *shape index* ( $m_S$ ) and *compactness* ( $m_C$ ) as it was expected, because the two shape measures are reciprocally defined. In the *orientation* category (M3) the last two measures are correlated, because both calculate an angle between the axes of the minimum bounding rectangle (MBR) of a building and the base line. The difference is that



measure M3-II is based on the main axis, while the measure M3-III considers both axes of the MBR (horizontal and vertical) and selects the smaller angle. In the *group* category (M4) the last two measures are highly correlated. A simple conclusion is to replace highly correlated measures of one category by a single measure, reducing the number of constraints.

**Table 5.2**  
Correlation values  
between pairs of  
measures. The grey  
cells highlight  
correlation values  
within the specific  
categories.

M1 – Size	Area	1	0.97	0.84	0.79	-0.62	-0.8	-0.03	-0.91	-0.13	0.07	-0.04	0.14	-0.07	-0.06	-0.24
	Perimeter	0.97	1	0.84	0.89	-0.71	-0.89	-0.05	-0.83	-0.09	0.08	-0.03	0.21	-0.02	-0.07	-0.23
	Length diam.	0.84	0.84	1	0.72	-0.72	-0.72	0.32	-0.75	-0.13	0.04	-0.05	0.12	-0.03	-0.04	-0.19
M2 – Shape	Shape index	0.79	0.89	0.72	1	-0.85	-1	-0.07	-0.59	-0.08	0.07	-0.02	0.32	0.02	-0.02	-0.15
	Build. concav.	-0.62	-0.71	-0.72	-0.85	1	0.85	-0.18	0.42	0.07	-0.15	-0.01	-0.3	0.11	0.06	0.09
	Compactness	-0.8	-0.89	-0.72	-1	0.85	1	0.08	0.6	0.08	-0.08	0.01	-0.3	-0.01	0.01	0.14
	Build. elong.	-0.03	-0.05	0.32	-0.07	-0.18	0.08	1	0.02	-0.01	0.08	0	-0.18	-0.14	-0.01	0.02
	Fract. dimen.	-0.91	-0.83	-0.75	-0.59	0.42	0.6	0.02	1	0.17	0	0.08	-0.03	0.11	-0.03	0.12
M3 – Orient.	M3 – I	-0.13	-0.09	-0.13	-0.08	0.07	0.08	-0.01	0.17	1	-0.16	-0.19	0.04	0.19	-0.04	0.01
	M3 – II	0.07	0.08	0.04	0.07	-0.15	-0.08	0.08	0	-0.16	1	0.76	-0.02	-0.07	-0.15	-0.18
	M3 – III	-0.04	-0.03	-0.05	-0.02	-0.01	0.01	0	0.08	-0.19	0.76	1	-0.21	-0.12	-0.2	-0.18
M4 – Group	M4 – I	0.14	0.21	0.12	0.32	-0.3	-0.3	-0.18	-0.03	0.04	-0.02	-0.21	1	0.33	-0.01	-0.05
	M4 – II	-0.07	-0.02	-0.03	0.02	0.11	-0.01	-0.14	0.11	0.19	-0.07	-0.12	0.33	1	-0.15	-0.25
	M4 – III	-0.06	-0.07	-0.04	-0.02	0.06	0.01	-0.01	-0.03	-0.04	-0.15	-0.2	-0.01	-0.15	1	0.84
	M4 – IV	-0.24	-0.23	-0.19	-0.15	0.09	0.14	0.02	0.12	0.01	-0.18	-0.18	-0.05	-0.25	0.84	1

Table 5.2 shows also strong correlation respectively anticorrelation between measures of different categories. For example the *shape index* is highly correlated with the *size* measures. One interpretation is, since the *shape index* measures the compactness of a building, that smaller buildings are more compact and therefore more simplified than larger buildings. A second interpretation is that in reality smaller buildings, such as houses, have a simpler, more compact shape than larger buildings like schools or hospitals. Further investigations are needed to evaluate whether such correlations depend on the map scale and are therefore caused by generalisation or not.

## 5.4 Detecting settlement types from building dataset

In this section the application of the PCA as a method for data reduction and visualisation is presented. The objective is the detection of settlement types from building data and the assignment of building attributes, that is, so called data enrichment (Ruas and Plazanet, 1996). Therefore, a set of artificial variables, called components, is obtained from the PCA and used for further data analysis and evaluation of settlement types.

Data Enrichment for generalisation purposes should equip the raw spatial data with additional information about objects and their relationships (Neun et al., 2004). That information is used for characterisation of map situations, conflict detection, and indi-



rectly also algorithm detection and evaluation. The assigned settlement types, obtained as a result of a structural analysis phase (Steiniger and Weibel, 2005a), can be exploited for building generalisation in a number of ways. They can be used for further types of data characterisation, e.g. restricting the search for building alignments by exclusion of rural and inner city areas, where buildings are too sparse or too dense, respectively. The settlement types can be used for algorithm selection, such that inner city buildings could be amalgamated instead of being displaced (for generalisation operations, see McMaster and Shea (1992)). In contrast too small buildings in rural areas would be enlarged instead of eliminated, since they may serve as landmarks for user orientation. Further, the settlement types could be used for the creation of small scale maps that depict settlement areas rather than individual buildings. Other kinds of usage could be imagined beyond the domain of map generalisation since knowledge about settlement types could help in applications as diverse as environmental and regional planning, traffic analysis for commuter train planning or spatial web searches (Heinzle et al., 2003, 2005).

#### 5.4.1 Method for extraction of settlement type regions in property space

The idea of the method is to describe types of settlement by a certain number of properties with respect to buildings. Using a building training data set the corresponding type regions are mapped into the so called property space (also called feature space in other branches of science). A similar approach is presented by Keyes and Winstanley (2001) for topographic feature classification using a property space defined by moment invariants and clustering techniques. In contrast our property space is a result of principal component analysis. If a building from another dataset is within an extracted settlement type region in property space, we can assign the settlement type to the building. Here one of the problems is to define the regions confining each settlement type in an  $n$ -dimensional property space. Thus, we like to reduce the number of properties without losing information on settlement type characteristics. In our case an optimal reduction would result in only two properties (dimensions), since this would alleviate the visual separation of the settlement types. In order to reduce the number of measurable properties we apply the PCA to obtain a set of transformed properties per building. By analysis of a number of maps and from previous research in building and settlement generalisation (Gaffuri and Trévisan, 2004) we identified five settlement types which are of interest. These are the three main types (1) urban area, (2) suburban area and (3) rural area and further the

two more specific types (4) commercial and industrial area and (5) inner city. The proposed method for the extraction of settlement type regions can be separated into 7 steps:

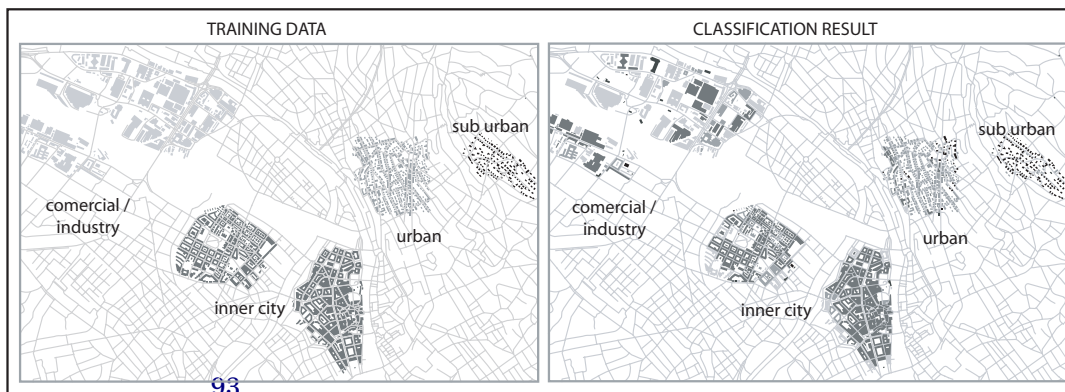
1. Characterisation of the five settlement types and definition of a number of measures.
2. Selection of at least one sample (training dataset) for each settlement type.
3. Characterisation of the buildings to obtain a set of property values for every building.
4. Data preparation for PCA (centring, standardisation, outlier detection and elimination)
5. Iterative process for settlement type separation:
  - (a) Application of PCA for reduction of properties, preferably to 2 or 3 transformed properties, the principal components.
  - (b) Adjustment of the number of input properties and the property values until a good settlement type separation is received (visual control). Here the number of input properties can be adjusted by using correlation analysis. The adjustment of the property values can be done by classification of values, resulting in a reduction of variance of a property.
6. Definition of the discriminating borders between the settlement types in 2 or 3 dimensional property / component space. This can be done either manually or using discriminant analysis methods (Duda et al., 2000). The result are regions in the transformed property space for the five settlement types.
7. Performing a test classification using the training data set to control the accuracy of the method.

Using the obtained parameters and regions for classification of other building data is based on the assumption that the training dataset is a representative subset of all building data. Based on that assumption and on acceptable results of test classification we save the settlement type regions and all the transformation parameters. From the PCA transformation we keep the so called component loadings, which describe the transformation from real building properties to the transformed properties. With these parameters we can now perform the settlement type detection on other building datasets without using the PCA anymore. Thus, we fix the mapping from the measured properties to transformed

properties and with it the obtained settlement type regions in property space. The following subsection will describe some issues of implementation and a test of the proposed method.

### 5.4.2 Realisation

A first classification into the five settlement types was made using topographic data (scale 1:25'000) of Zurich, Switzerland. We employed two kinds of settlement properties. On the one hand we used statistical properties of building groups (e.g. building density) and on the other hand we used properties of the individual buildings. [Steiniger and Weibel \(2005b\)](#) give an inventory of object properties and descriptive variables of inter-object relations.



Drawing from that a settlement type characterisation was made using the individual object variables: (1) building size (area), (2) building shape, (3) building squareness, (4) number of building corners, (5) building elongation and (6) the number of holes in a building polygon, which is an indicator of buildings in inner city areas. Further, two relational characteristics were identified: (7) the number of surrounding buildings and the ratio of buffer area to the building area inside the buffer. This last descriptive variable was evaluated by two measures, (8) one using a convex hull around all buildings in the buffer and another (9) using the buffer only. Thus, we have defined eight characteristics which are evaluated by nine measures. Consequently the property space, containing the settlement type regions, has nine dimensions.

After defining the set of measures a collection of sample sites, containing nearly 2000 buildings from the Zurich dataset, was established. Some of the sample data are shown in Figure 5.5 (left).

**Figure 5.5**

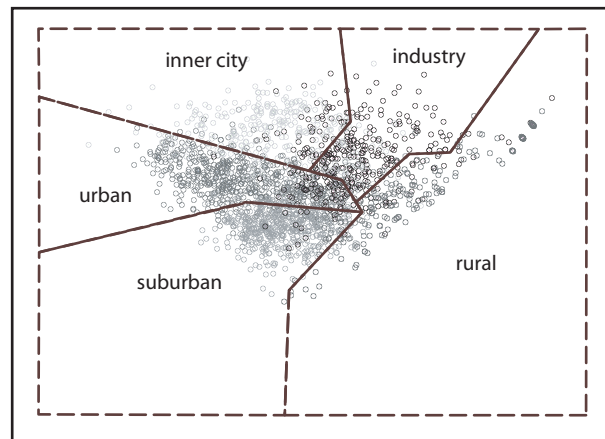
Part of the training data (left), where every group of buildings represent one settlement type. Classification results for the same data are shown in the picture on the right side. Correctly classified buildings have the same grey value as in the left picture. Data: VECTOR25, reproduced by permission of Swisstopo (BA057008).

According to steps three and four these buildings were characterised by use of the nine measures, and prepared for the PCA. This step and the following steps were realised by using the software package MATLAB for first experiments, and the open source GIS "Java Unified Mapping Platform - (JUMP)". Furthermore we used the open source packages JMAT and JMathPlot to realise the PCA and plot some results.

Data analysis of the variables prepared for the PCA showed that elongation and squareness discriminated the settlement types very poorly. In contrast, the relational measures (7, 8 and 9) showed good separation capabilities of the individual types. Further, a correlation analysis resulting from an initial PCA between the measures showed a medium correlation (value 0.5) between the shape index measure and the area measure. Thus, we decided not to use the measures building elongation and squareness to describe the settlement types. Now, a new PCA with only seven measures was performed.

A visual examination of a plot from the first two components reveals that a settlement type separation is hardly accomplished in a 2 dimensional component space. Further the evaluation of the explained variance per component showed that the first three components had variance values larger than one. According to the Kaiser criterion components with a value smaller than one are not necessarily needed to present the information contained in the dataset. From those two facts (visual examination and Kaiser Criterion) we concluded that at least three components are necessary for a sufficient separation of settlement types. The disadvantage of using a three component space is the more difficult visual detection of the type discriminating borders.

**Figure 5.6**  
Transformed and  
projected training  
data from 7  
dimensional property  
space (defined by the  
measures) onto an  
artificial property  
plane. The settlement  
type regions were  
defined manually.



Settlement type	No. buildings	No. build. correct classified	Probability matrix of classification				
			rural	industry/commercial	inner city	urban	suburban
rural	176	157	<b>0.89</b>	0.03	0	0.11	0.06
industry / commercial	365	193	0.11	<b>0.53</b>	0.16	0.17	0.04
inner city	316	218	0.02	0.13	<b>0.69</b>	0.16	0.01
urban	718	534	0.02	0.04	0.05	<b>0.74</b>	0.14
suburban	501	421	0.04	0.002	0.002	0.11	<b>0.84</b>

Therefore we searched for a heuristic way to reduce the three component space into two dimensions. Such a way has been found by classing the values for the Shape index (2) and the number of holes (6), which results in a reduction of variance. A new 3-dimensional plot of the first three components showed that a projection of the 3-D values on a plane is possible with the condition of settlement type distinction. Figure 5.6 shows the transformed (from 7 original measures to 3 components) and projected buildings from 3-D component space on an artificial property plane. Here the visualised settlement type borders and the type regions respectively, have been defined by hand for first tests on the usability of the proposed method. In future work it is planned to recognise the borders by use of classification methods.

**Table 5.3**

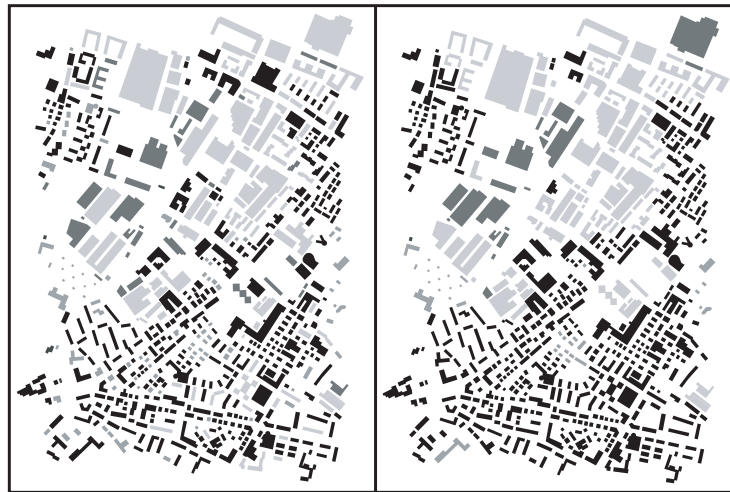
*Classification results for the training dataset. The dataset contains 2076 buildings from the Zurich study area.*

### 5.4.3 Results and discussion

The manually defined regions have been evaluated with the training dataset. The results portrayed in Table 5.3 and Figure 5.5 show good classification ratios for buildings in rural and suburban regions. A random classification for 5 classes would deliver an accuracy of around 20 percent. Also acceptable is the value of 74 percent of correctly classified buildings in urban areas and 69 percent in the inner city areas. Some problems occur for the detection of buildings in industrial areas. Here a third of the buildings has been identified as inner city or urban type. A reason is that the industry and commercial test sites were partly located near the city centre but also in the country side. Thereby the latter test sites, located in the country side, show features of urban areas. Thus, the borders between industry, inner city and urban buildings are fuzzily defined. This can be recognised as well in Figure 5.6 where the black dots, representing industry buildings, cover the two other regions as well. Sometimes one can find a building in a building group of another settlement type. For example an entrance building of an industry site or large buildings such as supermarkets located in a suburban housing area. Here the algorithm will assign the correct settlement type to the building, resulting in

an inhomogeneous picture. Thus, a spatial median filter may be applied to obtain more homogeneous regions (see Figure 5.7).

**Figure 5.7**  
Obtaining more homogeneous settlement type regions by application of a spatial median filter. Left picture: original classification; right picture: after application of the filter. Data: VECTOR25, reproduced by permission of Swisstopo (BA057008).



A further step of research on the settlement type detection approach will be the automatic recognition of the type borders in 2-D or 3-D property space. For this purpose several methods of classification techniques are possible. A first choice would be the use of linear discriminant analysis. Later machine learning techniques such as boosting (Schapire, 1999) should be applied to obtain potentially more accurate borders between the settlement types, similarly to the manually defined borders shown in Figure 5.6. Apart from a test with the training dataset two others, one Swiss and one North-European building dataset, have been tested as well. The visual evaluation of these results suggests a general validity of the method and the obtained parameters. An adaptation to countries with different settlement structure can be accomplished in two ways:

- A new fine tuning by adjustment of the settlement type regions in the property space/plane. This can be done either by discriminant analysis or by manual determination of discriminant borders. This would require a new sample dataset. The transformation parameters may stay the same.
- A new raw and fine tuning. Here new parameters for the property space reduction and the new settlement type regions in the property plane or component space have to be defined, using a country specific training dataset.



Opportunities for future work emerge from the exploitation of the detected types for analysis purposes or use in small scale maps. To that end, the creation of polygons from a group of buildings of one settlement type has to be investigated.

## 5.5 Detecting special buildings

A further application of the PCA and its data generalisation property is presented in this section. Here, we want to use PCA to find a priority to map generalisation for a number of extraordinary buildings. Such buildings might cause problems during an automated generalisation process, and should thus be treated separately (e.g. interactively). The approach to detect such buildings is the same as for the detection and elimination of outliers. However, our goal is not to eliminate, but rather to mark such outliers.

### 5.5.1 Method and experiment

The first step of the approach is to define a set of situations which could cause problems in building generalisation. Subsequently we describe such potentially problematic situations with a set of measures applied to buildings. In our case we did not focus on a specific generalisation situation but instead used a selection of the measures from the experiment on settlement type detection. The method to detect the buildings will now proceed as follows:

1. Define measures and apply them to the building dataset. In our experiment we used: area, shape, elongation, squareness, corners and the buffer measure no. 9, from 5.4.2. It should be noted that we did not use any orientation measure, since building orientation in European settlement data is often arbitrary. We believe that only a relative orientation measure with respect to surrounding buildings could be useful.
2. Data preparation (centring and standardisation) and computation of the PCA. Following the PCA it should be checked whether the correlation between the measures is low. Otherwise, redundant measures with high correlation should be excluded from further computation.
3. Calculation of mean values for every principal component.
4. Calculation of object distances in the component space from the centroid resulting from the mean values. The component variances are used to weight the distance components, since the distribution on the component axes is different.

5. Sorting by highest distance and selection of buildings with highest distance, representing the outliers and special buildings.

The described procedure is similar to Hotellings T2 test, a multivariate generalisation of Student's t-test. This test assumes a multivariate normal distribution and calculates normalised distances from the mean (Hotelling, 1931; Jackson, 1991; Zuendorf et al., 2003). The test is proposed for PCA outlier detection during computation of PCA in the software package MATLAB.

### 5.5.2 Evaluation of experimental results

Figure 5.8 shows a selection of special buildings, using one percent of all buildings with highest distance values (i.e. the top percentile) as a selection criterion. The result is as expected with respect to the used measures.

**Figure 5.8**  
Result of the  
identification of  
special buildings, for  
separate treatment  
during map  
generalisation. Data:  
VECTOR25,  
reproduced by  
permission of  
Swisstopo  
(BA057008).



The problematic point is how to define when a building is 'special'. Different methods are possible. One approach is to use a percentage value of all buildings, as we did. Here it can happen that there are really no unusual buildings are contained in the data set, but by using a percentage ratio we will always find some. Thus, defining a distance threshold might be more useful as a second approach. Then, every object with a distance value smaller than the threshold is an ordinary building. This approach needs a fixed transformation used for all building data sets since adding or removing one building to/from the data to analyse may change the variance of the data, with it the PCA transformation parameters and eventually the distances. The idea of the latter approach is similar to our approach of the settlement type detection example.



A third approach, however not yet tested, might be to search for breaks in the histogram of all distances. Finally further experiments on this topic will have to be carried out since the presented results should only hint at the applicability of the method.

## 5.6 Conclusion

The paper describes the usage of Principal Component Analysis in the process of automated generalisation. Several applications have been identified, for example the settlement type classification and the detection of special buildings as part of data analysis in the generalisation process. PCA is also applied for the investigation of measures and constraints, which are needed for conflict analysis and evaluation as well as the selection of generalisation operators. To allow a more formalised treatment of constraints a model called constraint space was introduced, which has been derived from generalisation state diagrams. The constraint space can be transformed with respect to a representative test data set. Therefore correlations between constraints have to be identified and in conclusion the number of constraints can be reduced.

When generalising building alignments the homogeneity of the groups should be preserved. Therefore several measures evaluating the homogeneity of groups have been analysed with PCA. High correlation values between measures appeared when the same cartographic phenomenon is described, as is the case for the measures area and perimeter representing the size of an object. Hence, the two measures can be replaced by only one of them. However, high correlation values also occur if a relation exists between different geographic phenomena, e.g. on the average smaller buildings are more compact than bigger buildings. Such relations would be kept or emphasised with respect to map purpose. The paper presents mainly investigations of shape, size, orientation and group measures, because they are the basis of our homogeneity constraint definition. Further research will have to analyse also the influence of generalisation operations on constraints, in particular correlations of constraint value changes. These correlations may provide information on side effects of generalisation operators; for example, a generalisation operator that is applied to solve a particular constraint may cause other constraint violations.



## 6 Multi-Representation Databases with Explicitly Modelled Horizontal, Vertical and Update Relations

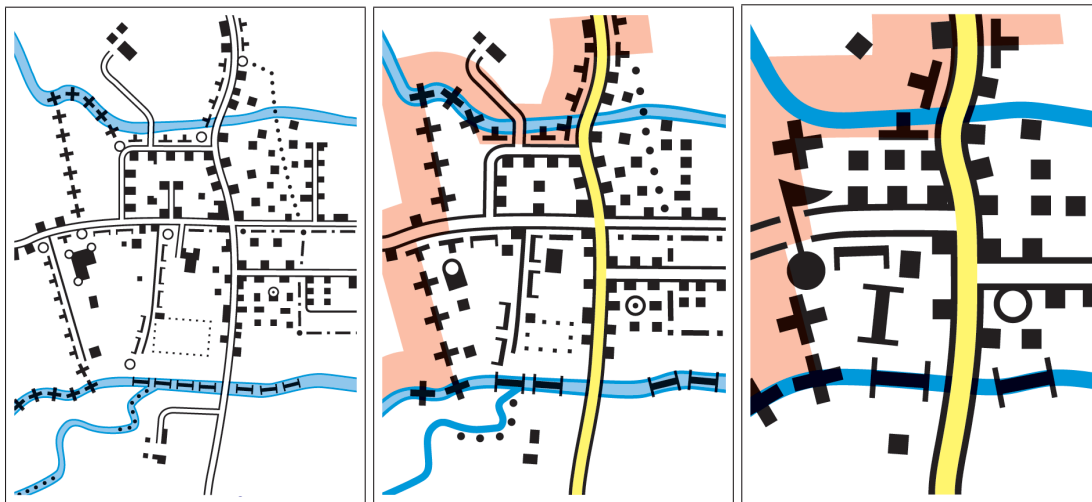
Matthias Bobzien, Dirk Burghardt, Ingo Petzold,  
Moritz Neun and Robert Weibel

published in  
Cartography and Geographic Information Science, 2007

**Abstract:** This paper presents a new approach combining multi-representation databases with the generalization and update process. It leads to a tightly integrated model, which is a part of the existing GIS cartographic solution *axpand*. The approach is based on the mathematical concept of relations and, in particular, on three different types of relations: horizontal (within one resolution), vertical (between different resolutions) and update. Horizontal relations allow the representation of relationships between features within one resolution. Examples are partitions, neighborhoods and topology. The vertical relation represents the relationship between features of different resolutions. This originates from a generalization or matching process. The update relation describes temporal changes of features. After a detailed theory about the relation types introduced, the paper continues with a discussion of their similarities and differences with a focus on implementation in a multi-representation database. A prototype demonstrates the generalization of buildings and roads from vector data at 1: 25 000 to 1: 50 000 scales. The paper ends with conclusions and an outlook on further research tasks.

## 6.1 Introduction

The lack of fully automatic real-time generalization functionality has forced research and development of multi-representation databases (MRDB). A wide range of applications for MRDB exists in conventional map production as well as in web cartography (Jones et al., 1996; Devogele et al., 1996; Hampe et al., 2003). These have, however, been realized in only a few cases. The aim of this paper is the presentation of the implementation of an MRDB in a cartographic map production system. Research and development focuses thereby on explicitly modeled relationships (Neun et al., 2004; Neun and Steiniger, 2005). These support the generalization process in two ways: First, by aiding in the analysis of data of one resolution and thus supporting the process of generalization, and second, by maintaining incremental updates.



**Figure 6.1**

Example of generalized map features in different resolutions (scales 1 : 25 000, 1 : 50 000 and 1 : 100 000) and corresponding symbolization. Figures taken from SSC (2005), with permission from the Swiss Society of Cartography.

The research and development work corresponds to the current requirements of various European national mapping agencies. Their aim is to introduce flexibility and process consolidation by putting in place a GIS-based cartographic production line in conjunction with MRDB, containing linked features at different resolutions. A typical map example is shown in Figure 6.1. It depicts a resolution change between three different resolutions. The modeling discussed here is intended to cover any range of resolution change.

The explicitly modeled relationships are divided into three different types: *horizontal*, *vertical* and *update* relationships. Horizontal relationships describe the relations between features in one

resolution (level of detail, LOD). Examples are partitions, neighborhood and topology, semantic, structure and patterns (Duchêne, 2004; Steiniger and Weibel, 2007). Vertical relationships connect features that represent the same real-world entities. These relations between features of different resolutions are created during a generalization process or by matching operations. The update relationships describe temporal changes of features that may be derived from an updating process or from matching of datasets with different time states. For each of these types of relationships we present existing approaches for their creation and analyze their dependencies.

The connections between the three types of relationships are explained briefly: *Horizontal relations* support the automated generalization process. The result of this process is stored in different resolutions, connected by a *vertical relation*. This relation is required for automated incremental updating which is triggered by updates, and is modeled through an *update relation* between different temporal states. Furthermore, the update relation enables spatio-temporal analyzes. The choice of the terms "horizontal" and "vertical" refers to the picture of resolutions as a stack of data layers of different resolutions (Neun and Steiniger, 2005; Steiniger and Weibel, 2007). Horizontal relations affect only a single layer (or resolution), while the vertical relation extends across the stack of (resolution) layers. The terms are not meant in a geometrical (three-dimensional) sense.

The main result of our work is the integration of the three types of relationships in one common model within a MRDB. The similarities and differences between horizontal, vertical and update relations are examined. An approach for explicit representation of these relationships is proposed. The result is an enriched MRDB. The advantages of this enriched MRDB are the support of the generalization process, improved possibilities of data analysis through resolution and time, as well as the consistent management of geographic and cartographic data. The results can be used for exchange with generalization services as well.

## 6.2 Theory and definitions of horizontal, vertical and update relations

Relations are a concept widely used in a variety of fields. They are all based on the mathematical concept of relation in which a relation is defined as a set of tuples with a fixed length, each tuple being built from given sets. Formally a relation  $R$  is a subset of the Cartesian product of a couple of sets  $A_1, \dots, A_n$ , thus  $R \subseteq A_1 \times \dots \times A_n$ . The number  $n$  is called the *order* or *arity* of the

relation. In many cases only relations of cardinality 2 are considered, expressed by "a is related to b". In a broader sense, relations of any order  $n$  with  $n \geq 1$  can be considered and the relation is called  $n$ -ary relation. The usual case of  $n = 2$  is called a binary relation, whereas the cases  $n = 1$  and  $n = 3$  are called unary and ternary relations respectively. Unary relations are a special case, in the sense that a relation is simply a subset of the given set  $A_1$ .

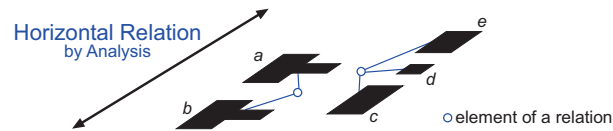
The following sections introduce the types of relations which are important for analysis, generalization, handling and updating of geodata, namely horizontal relations, vertical relation and update relations. Subsequently, combinations of these types will be examined and the possibilities of describing *relations of relations* will be discussed.

### 6.2.1 Horizontal relations

Horizontal relations characterize map features of one specific resolution or level of detail (LOD) on a defined time stamp. Examples are partonomic relations, neighborhood relations, structural relations or patterns, semantic relations and hierarchical relations. The order of the horizontal relations can be between  $1 \dots n$  depending on the number of features, which are characterized through the relation. A special case is a horizontal relation of order one. An example for this case is the modeling of partitions through horizontal relations, whereby one single feature can create one partition. Another unary horizontal relation of a semantic nature, modeled practically in every GIS, is the assignment of a feature to a specific class, for example *Mainstreet* to the class road.

Very common is the modeling of binary relations applied to groups of features. For example a building alignment is modeled as a meso object (Ruas, 2000; Ruas and Holzapfel, 2003; Li et al., 2004). The relations between the individual building features and the alignment of a meso object are binary part-of relations. Our definitions deviate from this explicit introduction of groups, or meso objects, and reduce the model to individual features and their relations only. Figure 6.2 shows a schematic view of two horizontal relations represented by the blue circles. Two building alignments are modeled as a relation, consisting of two elements, of which one is the tuple (a, b) and the other the tuple (c, d, e).

Horizontal relations can be differentiated according to the computation complexity and the frequency of their usage. The range reaches from less complex and single usage, such as feature attribute comparison, to complex and repeated usage, such as alignments. Depending on these different cases, the relations are calculated on-the-fly, sometimes repeatedly, or calculated once and

**Figure 6.2**

*Horizontal relations on the example of building alignments: The horizontal relation has two elements, one connecting the buildings a and b, and the other connecting the buildings c, d and e.*

stored persistently. To allow a homogenous access to horizontal relations, in all cases the same data structure is used (see section about Prototype and Results). Advantages become obvious if horizontal relations are interpreted as sets of features with certain characteristics. The application of a common structure allows the usage of set operations like combinations and intersections of related features which lead to new horizontal relations. Furthermore, a common structure allows the construction of functions and methods that receive a relation element as a parameter or give a relation element as a result of a calculation. Even standardization for the use in web services is conceivable (Neun et al., 2006).

The number of possible horizontal relations is tremendous, and therefore a decision must be made as to which horizontal relation to calculate. This decision depends on the application. Of even higher importance is the decision of whether to store the relation persistently or to calculate it on-the-fly. This decision must be made for each relation separately. We recommend the first for complex and frequently used relations, the latter for simple relations and relations that are seldom used. In the following section, horizontal relations are presented with a focus on map generalization of topographic maps.

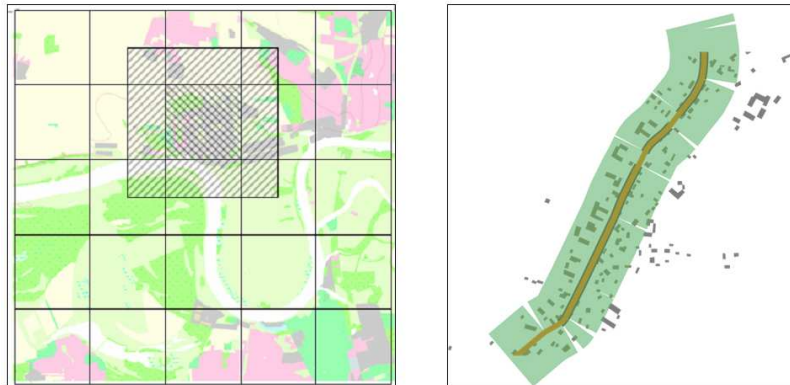
### Partonomic relations (partitions, global neighborhood)

Partitions subdivide the map space in such a way that generalization tasks can process them independently. The aim is to identify areas in which to restrict the influence of generalization operators. The simplest case is a grid-based subdivision of the map space using a buffer area around each generalization zone (Figure 6.3, left). The features inside the buffer area may influence the generalization of features inside the grid cells, but are not modifiable themselves. Exceptions are applied to area or line features in the case when the biggest part is situated inside the generalization zone. The boundary problem can be handled with this technique.

Besides these regular subdivisions of map space, contextual subdivisions are possible as well. An example is the trans-hydrograph proposed by Timpf (1998) that describes a graph structure derived from transportation and hydrology networks. The trans-hydrograph can be used to subdivide the task of building gener-

alization since buildings must stay inside the faces of the trans-hydro-graph.

**Figure 6.3**  
Examples of  
horizontal relations -  
buffer area around a  
generalization zone  
(left) and line  
buffering (right).



Explicit storage of partitions allows the definition of generalization zones that describe regions, in which the same parameter values of generalization operators have to be applied. The generalization zones are dependent on the object classes and the applied operators.

#### Neighborhood relations (distance based, topological, graph based)

In contrast to space partitioning approaches, individual neighborhoods can be modeled as a relation in a way that each feature holds information about its local surrounding. A simple approach is the distance-based storage of features using point or line buffering (Figure 6.3, right). During position or shape change only the related features have to be considered. The linkage may work in both directions, which additionally informs the surrounding features if position or shape has changed.

Another way of modeling neighborhood relations is the application of graph structures. Examples are topological data structures, transport and neighborhood graphs, triangulations or surface networks. Here more than the pure distance information has to be considered. Topological structures especially support the modeling of spatial relations like *connectivity*, *overlap*, *inside* etc. (Egenhofer and Franzosa, 1991; Egenhofer and Herring, 1991). Other neighborhood relations that can be derived from graph structures are *orders of neighborhood*, *relative orientation* or *relative size*.



### Semantic relations

Semantic relations define the classification of features. In current GIS or map production systems these semantic relations are explicitly available through the data model and the layer structure. In contrast to other horizontal relations, they must be defined by the user and cannot be calculated from the data in general. Semantic relations are a prominent example of a unary relation, each relation describing a subset of all features that belong to a certain feature class.

Semantic relations also determine the selection and orchestration of generalization operators, either interactively through the cartographer or through the definition of different cartographic constraints for an automated process control. The selected generalization operators must correspond with the map feature class. For instance, the *building simplification operator* can only be applied to features of the class building, or, features belonging to a railway network will be generalized with different operators or parameters than features belonging to a river network.

Finally semantic relations have an influence on other horizontal relations: rivers, for example, must not cross one another and terrain influences the generalization of rivers and roads. Therefore horizontal relations must be defined with respect to the specific feature classes.

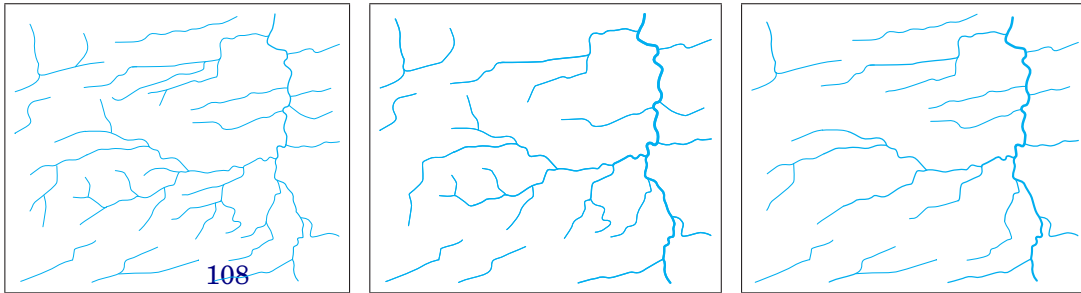
### Patterns or structural relations

The human eye reacts quite sensitively to patterns, repetition, density differences, parallelism, orientations or global trends, which means they have to be modeled and made available during the generalization process. Most prominent relations are alignments, especially for buildings, which have been investigated quite extensively [Christophe and Ruas \(2002\)](#); [Ruas and Holzapfel \(2003\)](#); [Li et al. \(2004\)](#). Another example of structural relations can be found within river networks (Figure 6.4). Here they allow the distinction between meander, canal-like, tree structures or tectonic based parallel orientation.

Further examples can be found for other feature classes such as strings of islands, street patterns ([Heinzle et al., 2005](#)), vegetation, terrain or settlement patterns.

#### 6.2.2 Vertical relations

The vertical relation links features that represent the same real-world phenomenon in different map resolutions or levels of detail (LOD). Thus a real-world feature can have several representations

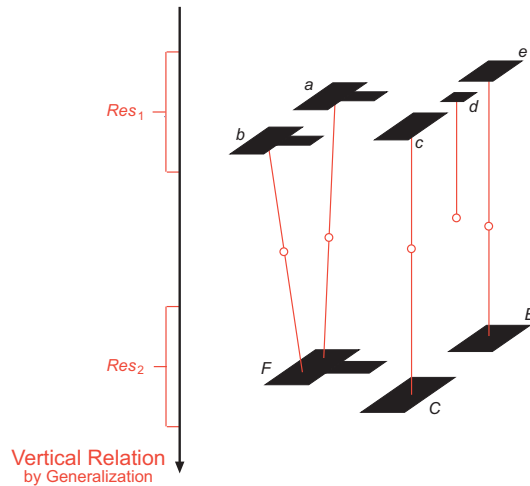
**Figure 6.4**

*The detail selection of a river network (original, left) without consideration of structural relations (middle) and with explicit focus on dominant orientation (right). Figures taken from SSC (2005), with permission from Swiss Society of Cartography.*

in a MRDB. In this paper, the term *feature* denotes the *representation* of a real-world feature in a MRDB. Traditionally, linkage is created by storing a unique identification value (ID) from the corresponding features of the connected resolutions. In our approach, the vertical relation is made explicit through the creation of instances of the vertical relation class, whereby additionally relevant meta information about the production process can be stored (see section about Similarities and Differences). Examples of such information are applied generalization operators together with their parameters that produced the derived features in the case of generation by generalization, and the matching operator in the case of generation through the matching process.

The vertical relation is a crucial characteristic of MRDB. It supports, for instance, update propagation, quality and consistency checks and multi-resolution analysis functions as well as zoom functionalities for web mapping and mobile applications (adaptive zooming). If the features of a lower resolution were derived by cartographic generalization, the vertical relation may store information about the applied operators and parameters. If existing datasets from different sources are integrated, the links have to be created by means of automatic matching techniques. Thereby additional information about the matching probability can be stored.

Figure 6.5 shows a schematic view of the vertical relation: The buildings *a* and *b* are aggregated into building *F*; buildings *c* and *e* are generalized into *C* and *E*, while building *d* is omitted during the generalization. Therefore, the relation consists of five elements, consisting of four 2-tuples (*a*, *F*), (*b*, *F*), (*c*, *C*), (*e*, *E*) and one 1-tuple (*d*). For discussion of cardinality see for example (Bobzien et al., 2005b).



**Figure 6.5**  
Schematic view on  
the vertical relation.

### 6.2.3 Update relations

The update relation links features representing the same real-world phenomenon in different time states. The time states represent sequential updates of geodata. Spatio-temporal GIS could be applied for a realization of the proposed update relations.

Update relations support versioning and temporal analysis. In contrast to the vertical relation, the update relation stays in one resolution. The update relation might be derived from an incremental updating process. The term incremental updating (Kilpeläinen and Sarjakoski, 1995; Haunert and Sester, 2005) describes the propagation of updates within a base data set to another data set, whereby the generalization deals primarily with the features that have changed over time (Harrie and Hellström, 1999). Additionally, neighboring features (characterized through horizontal relations) have to be considered as well, since they influence the generalization process. These neighboring features may be identified by horizontal relations (see above).

Three different update operations exist: *insert*, *remove* and *change*. In the first two cases, unary update relations, consisting of the set of inserted or removed features, are created respectively. The third case is described by a binary or *n*-ary relation of changed features. These may be divided into geometric and semantic changes (see below).

Update operations may initialize an incremental generalization (Haunert et al., 2006). The creation of an update relation may have an effect on vertical and horizontal relations. These relations have to be checked and possibly renewed. The interactions

between horizontal, vertical and update relations are described in more detail in the next section.

An additional challenge in cartographic generalization, especially for incremental generalization, is the preservation of manual edits. Manual corrections will always be necessary because a fully automated generalization solution that works in every case does not exist and will not exist in the near future (Schuurman, 1999). The process of manual correction should be kept as accurate as possible, because repetition is time consuming and therefore expensive. A cost-efficient solution can be offered by a system that only re-generalizes in required cases as a result of change in the source data. Horizontal and vertical relations allow the identification of these tasks.

**Figure 6.6**  
Schematic view of an  
update relation.

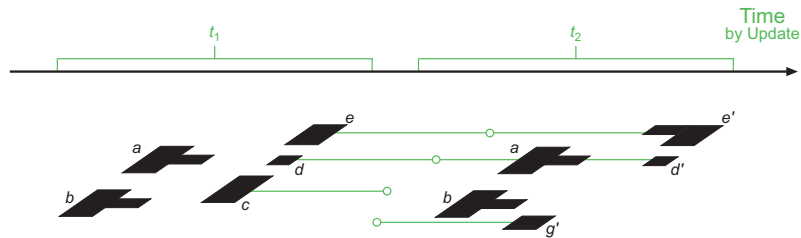


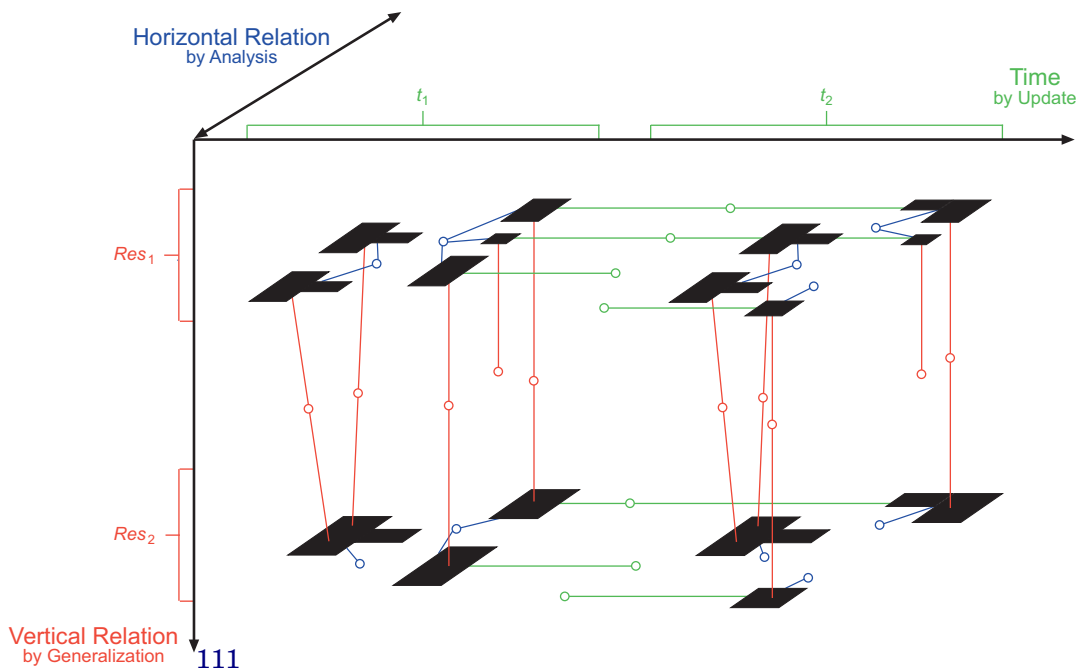
Figure 6.6 shows a schematic view of the update relation:  $t_1$  and  $t_2$  denoting two different time states. Features  $a$  and  $b$  remain unchanged. Feature  $e$  undergoes a geometric change into feature  $e'$ , while feature  $d$  undergoes a semantic change into feature  $d'$ . Feature  $c$  is deleted, while feature  $g'$  is inserted. Further examples could include aggregation and splitting of features (see section about Similarities and Differences), thus the update relation must provide an  $m : n$  cardinality.

#### 6.2.4 Combination of Horizontal, Vertical and Update Relations

Features can be characterized through one or more instances of the above introduced relation types. Figure 6.7 shows a combined example of horizontal (blue), vertical (red) and update relations (green). Therefore the same situation is presented four times, at two different time stamps ( $t_1, t_2$ ) and in two different resolutions ( $Res_1, Res_2$ ). The initial state consists of five buildings aligned into two rows shown at the upper left. The two alignments are modeled by horizontal relations (blue). The situation after generalization is shown at the lower left of the figure. The buildings are typified

and enlarged and the derived buildings are connected to the corresponding ungeneralized buildings by the vertical relation (red).

The situation at a second time stamp is visualized in the upper right of the figure. The three possible update operations *delete*, *insert* or *change* are applied. The corresponding update relations are shown in green. The update that was propagated to the second resolution through an incremental updating is depicted in the lower right of the figure. Therefore, either the updated situation of  $Res_1$  requires a completely new generalization, or only features that were updated will be generalized (Cooper, 2003; Jahard et al., 2003). The second case implies that for those features which were not updated the corresponding generalized features of  $Res_2$  at  $t_1$  do not need an update to  $t_2$ . However, if the surrounding features have changed, this may influence the generalized feature, although the feature itself was not updated. Thus the horizontal relations (describing the surroundings) have to be considered as well.



**Figure 6.7**  
Combination of  
horizontal relation  
(blue), vertical  
relation (red) and  
update relation  
(green).

The example shows the interference of the three types of relations, which are all needed for the process of automated incremental updating.

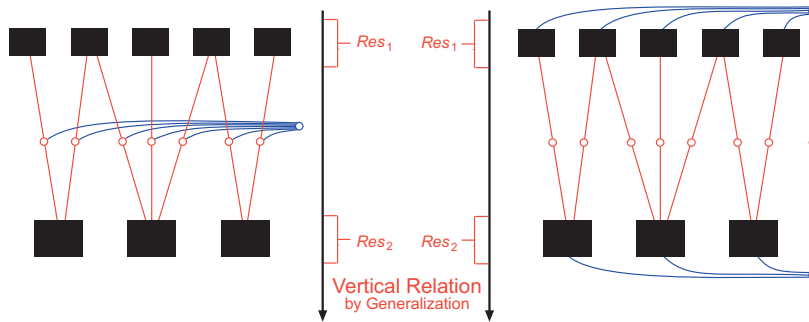
### 6.2.5 Relations combining relations

An interesting question refers to the interrelationship between relations of different types. First the interrelationship between horizontal and vertical relations will be discussed, second, the relationship between horizontal and update relations.

Horizontal relations characterize important contextual information within one resolution. If the horizontal relations are modeled within the origin and target dataset, they can be used to measure the quality of the generalization process. Preservation, grade of change and loss of horizontal relations lead to criteria for quality measurement. There are two possibilities for making this happen.

The first possibility considers the fact that several generalization operators consider horizontal relations during execution. This implies that the resulting vertical relation contains the information of the horizontal relations. Figure 6.8 (left) depicts a situation of five buildings that form an alignment in the origin dataset and are generalized into three buildings in the target data set. The information about the alignment is stored implicitly within the vertical relation, which, in turn, was created by the generalization operator that took the horizontal relation into account.

**Figure 6.8**  
Combination of horizontal relations and vertical relation.  
Left: The vertical relation elements are connected by a horizontal relation.  
Right: The building alignments (horizontal relations, blue) in different resolutions are connected by a vertical relation (red).



The second possibility is intended for those cases in which a generalization operator acts independently of horizontal relations. The horizontal relations of the target dataset have to be built up after the process of generalization. An explicit modeling of the vertical relation between the two horizontal relations of each dataset, origin and target, can now take place. This relation may be used for a subsequent quality analysis of the performed generalization operator. Figure 6.8 (right) depicts the situation of an explicit modeling of the two building alignments in both resolutions. The vertical relation connects the buildings of both resolutions and, additionally, the alignments are connected by a vertical relation.

Updates may change the context of a situation and thus affect horizontal relations. This may lead to deletion, change or creation of new horizontal relation elements after an update. To react to these changes-especially in the context of automated incremental updates-it is possible to explicitly model an update relation between horizontal relations. Depending on the change of the horizontal relation, a specific generalization update strategy may be selected. Different strategies for incremental update are discussed in [Bobzien et al. \(2005b\)](#). Either information gathered in a prior generalization process may be reused, or the generalization has to be re-calculated.

## 6.3 Similarities and Differences

In this section we discuss similarities and differences of horizontal, vertical and update relations. Common to all three relations is the concept of data enrichment in a standardized way, through a common class *Relation*. The specification of the class *Relation* enables the explicit storage of meta information, relevant for generalization. The main differences refer to the arity of the relations and to the fact that horizontal relations describe *states* of features whereas vertical and update relations describe the *change* of features in a MRDB.

### 6.3.1 Data Enrichment

All three relation types intend to enrich feature models by supplementary characterization of interrelationships through space, semantics, resolution and time. The relations characterize features and describe their association by meta information, supporting the automated generalization process. In detail, the meta information relevant for generalization will improve the quality, efficiency and the grade of automation of the generalization process.

Data enrichment comprises two aspects: the first one is the analysis of implicit relations and its explicit modeling. Examples are the modeling of neighborhood relations via topology, the matching of two datasets ([Sester et al., 1998](#)) and the modeling of context information ([Mustière and Moulin, 2002](#)). The second aspect is the enrichment of these relations by information that is not contained in the datasets, such as process information. Horizontal relations describe states of features whereas vertical and update relations describe changes of features. Therefore the process information is only important for the latter cases.

### 6.3.2 Detailed Specification of Relations

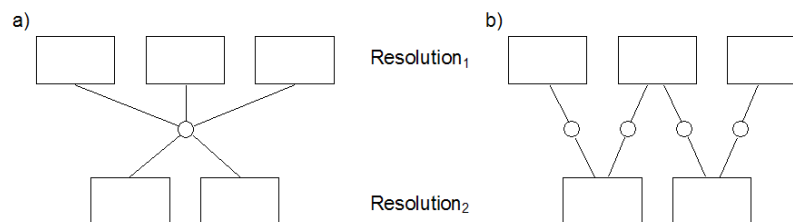
All three relation types can be specified in more detail. Specializations of horizontal relations are, for example, alignment relation and partitioning relation (see above). The vertical relation may be divided into matching relation and generalization relation. A matching relation is built by a matching algorithm or by manual matching. The relation can hold additional information, e.g. probability values. The generalization relation is built by a generalization operator. Additional information can be stored, such as parameters and involved features. The update relation can be distinguished by the type of operation into the insertion relation, deletion relation and change relation. The change relation can be enriched by information on the kind of changes, like geometric or attributive changes. All kinds of update relations contain a time stamp and additional information about the update process.

### 6.3.3 Arity of Relations

All relations can be modeled as  $n$ -ary relations. This modeling is suitable for horizontal relations since the number of related features can vary. For instance a specific partition as an element of the partition relation can contain an arbitrary number of features.

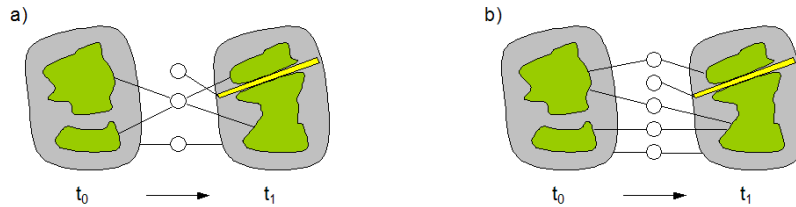
For the vertical relation and the update relation the modeling as an  $n$ -ary relation is possible as well. However, binary relations are more suitable here since they imply more information than an  $n$ -ary relation. If an  $n$ -ary modeling is chosen, it is uncertain which feature emerged from which other feature. Two examples in Figure 6.9 and Figure 6.10 confirm this statement for the vertical relation and update relation respectively.

**Figure 6.9**  
Two different  
modelings of a  
vertical relation:  
modeling by an  $n$ -ary  
relation (left) and by  
a binary relation  
(right).



In Figure 6.9 (a) the typification of three buildings into two buildings is modeled by an  $n$ -ary relation. If the leftmost or the rightmost building of resolution 1 receives a minor update both buildings of resolution 2 have to be updated. In contrast the binary modeling in Figure 6.9 (b) enables a direct detection of the building that has to be updated.





**Figure 6.10**  
Two different  
modelings of an  
update relation:  
modeling by an  $n$ -ary  
relation (left) and by  
a binary relation  
(right).

Binary relations are also more suitable for the modeling of the update relation as shown in Figure 6.10. The original dataset of time stamp  $t_0$  consists of two forest areas. At timestamp  $t_1$  a road was added and the forest area was enlarged. In Figure 6.10 (a) all four forest areas are related to one update relation by an  $n$ -ary relation. The binary relation in Figure 6.10 (b) is more specialized, which has advantages for the calculation of incremental generalization. The figure also shows an example of a unary update relation characterizing the inserted road feature.

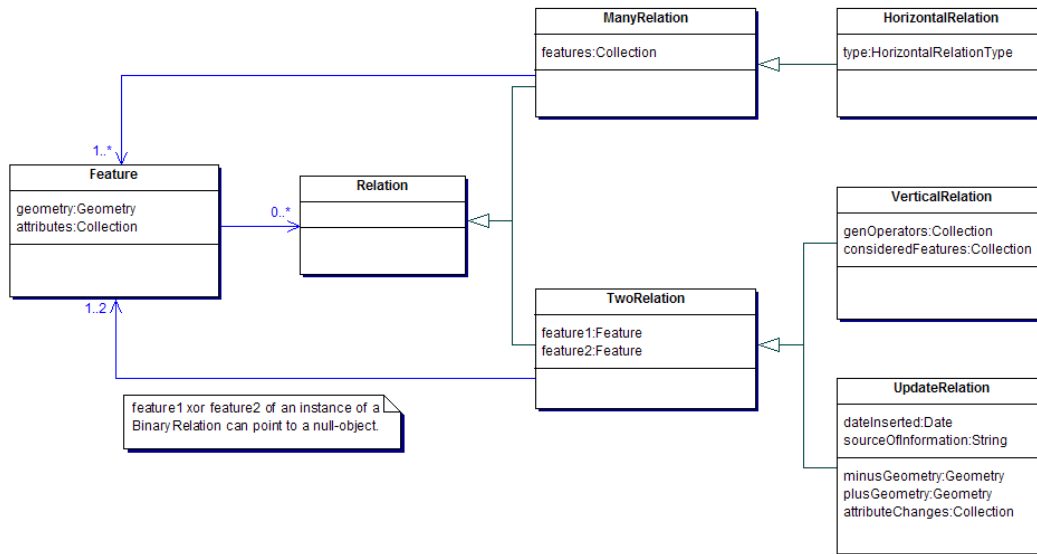
#### 6.3.4 Object-oriented modeling of relations

In this subsection an object-oriented modeling of the relations is presented. The standardization of the relations enables the application of operations to any relation without distinction between the specific types. A common operator may include all features that are characterized by the given relation element. Additionally, standardization allows the exchange of relations in a homogenous manner, e.g. via (web) services.

The object-oriented model is depicted in Figure 11. Each relation is represented by one of the following classes: HorizontalRelation, VerticalRelation and UpdateRelation. Depending on the multiplicities, these relations can be differentiated as shown in UML-Diagram. The diagram is reduced to the main classes and main attributes.

Each feature may have none, one or many associations with the class Relation. The relation can be specialized to the classes NArYRelation and BinaryRelation. These subclasses differ in the multiplicity of their associations to the class Feature.

BinaryRelation is associated with one or two features. It can be specialized to the classes VerticalRelation and UpdateRelation. The class VerticalRelation stores the generalization operator as well as all features that influence the generalization. The class UpdateRelation holds the information about the date of update, source of the information, differences of the geometry and modifications of the attributes.



**Figure 6.11**  
UML-Diagram of the implementation of relations, reduced to the main classes and attributes. An open arrow represents a directed association, with given multiplicity. A closed arrow represents a specialization.

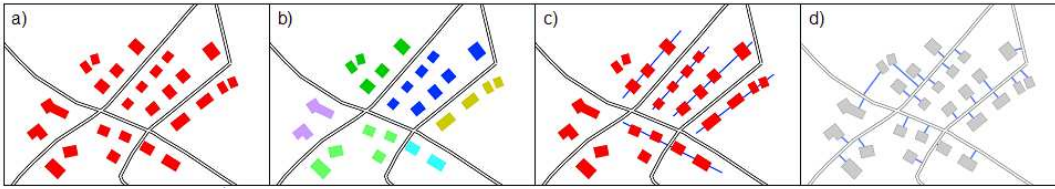
An `NaryRelation` is associated with one or many features. The class `NaryRelation` has one subclass `HorizontalRelation`. The main attribute of the `HorizontalRelation` is the type of the relation. This attribute describes the kind of the horizontal relation, e.g. alignment, partitioning or feature class membership.

## 6.4 Prototype and Results

A prototype that demonstrates the applicability of the theoretical concept has been implemented in the cartographic GIS *ax-pand* system (Axes-Systems, <http://www.axes-systems.com>). At the time of this writing, three horizontal relations as well as the vertical and the update relation are included in the prototype.

Implementation is object-oriented using the programming language Java. The Java Topology Suite (JTS) is used for geometrical calculations and modeling. The visualizations in this section are created on the Jump Unified Mapping Platform (JUMP). The data used for the following examples are taken from a typical VECTOR25 dataset and show a small section consisting of buildings and roads.

The original situation is depicted in Figure 6.12 (a). Screenshots (b)-(d) show three different horizontal relations that have

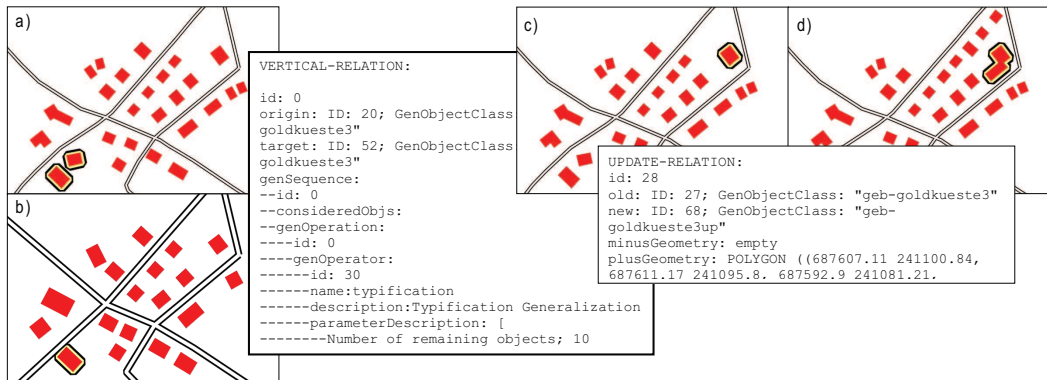
**Figure 6.12**

*Horizontal Relations:*  
 (a) original situation  
 (b) partitioning by  
 trans-hydro-graph (c)  
 building alignments  
 (d) neighborhood  
 relation between  
 buildings and nearest  
 road.

already been described in this paper. The relations are derived automatically from the original dataset. The partitioning of the buildings is calculated using the trans-hydro-graph (b). Building alignments are detected by consideration of several constraints such as size, shape, orientation and distances between buildings and towards roads (c) (Burghardt and Steiniger, 2005). The neighborhood between buildings and roads is computed by search of the nearest road to each building (d). The relations are modeled by the class `HorizontalRelation` as described in Figure 6.11. Partitioning and alignments are  $n$ -ary relations while the neighborhood relation is modeled as a binary one. All three horizontal relations shown are used for the generalization process as shown below.

The left part of Figure 6.13 shows an example of the vertical relation. Figure 6.13 (a) presents the original situation (the same as in Figure 6.12) while Figure 6.13 (b) shows a generalized situation of scale 1 : 50 000 which is derived automatically. The applied generalization operators consider the symbol change, e.g. buildings are displaced from the streets. The highlighted buildings in (a) are typified into the highlighted building in (b). The text to the right of this example is the alpha-numerical output of the two highlighted elements of the vertical relation. The generalization of all features was calculated using horizontal relations, namely partitioning and alignments for typification, and neighborhood for displacement. The horizontal relation is modeled by the class `HorizontalRelation` as described in Figure 6.11. Information about the performed generalization process is stored within. This is a binary relation as visualized in Figure 6.5. The vertical relation is the core of a MRDB and used for automatic incremental updating as shown below.

An example of the update relation is shown in Figure 6.13 (c) and (d). To the right several updates have taken place: One building is deleted (upper middle), three buildings are inserted (upper right) and the geometry of the highlighted building is changed. The alpha-numerical output below describes the highlighted element of the update relation. In the ungeneralized dataset the update can be performed interactively by deleting, inserting or up-

**Figure 6.13**

Vertical relation between features of different resolutions with scale dependent symbolization (left) and update relation between features in one resolution (right).

dating features. Alternatively, this process can be carried out by utilizing an update dataset. In the generalized dataset the update relation is created only by automatic incremental updates as shown in Figure 6.14 and discussed below. The update relation is modeled by the class `updateRelation`. This binary relation is depicted in Figure 6.6. The update relation realizes a multi-representation in the time dimension, and thus gives a different view on MRDB. The explicit modeling of the update relation is necessary for automatic incremental updating as well as for history management, versioning and spatio-temporal analysis.

**Figure 6.14**

Interference of horizontal, vertical and update relation. Situations: (a) original, (b) updated, (c) generalized original, (d) updated and generalized. Compare Figure 6.12 and Figure 6.13. The order of the screenshots is analogue to the order in Figure 6.7.

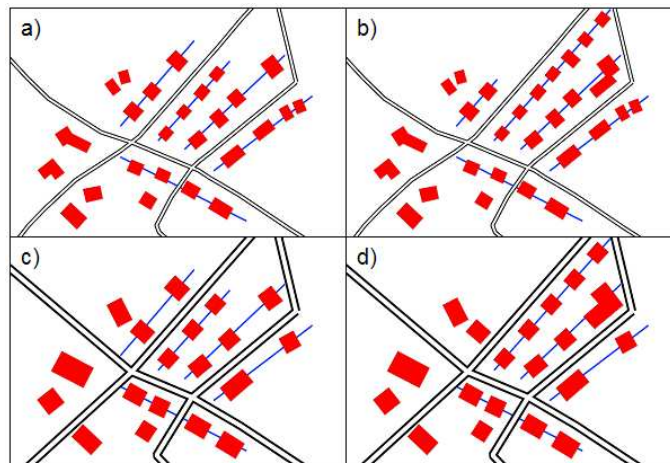


Figure 6.14 is analogue to Figure 6.7 and shows an example of an automated incremental updating. In all four screenshots

the horizontal relation of alignments is shown. Figure 6.14 (a) depicts the original situation, (b) the updated situation, (c) the generalized situation and (d) the result of the incremental updating process. The preservation of horizontal relations is discussed above. In the example, the horizontal relation alignment has to be preserved during generalization. During update, the horizontal relations may change as shown in the example in the two upper alignments: one is shortened, the other elongated. The change of the horizontal relation during update influences the incremental update directly. Quite often the generalization of an updated feature is identical to the generalization of the original (non-updated) feature. In this case no update occurs in the generalized dataset and therefore no update relation in the generalized dataset is modeled.

## 6.5 Conclusion

This paper presents a framework for data enrichment through common modeling of relations in a multi-representation database. The relations are defined as further characterizations of features and their changes over resolution and time, with focus on cartographic generalization.

Horizontal relations deliver input information for generalization operators in a standardized way to improve the quality of the generalization results and increase the calculation efficiency. They support the modeling of groups of features, similarly to the concept of meso objects. The vertical relations hold information about generalization operations or matching procedures to support, for instance, update propagation, quality and consistency checks, multi-resolution analysis functions, as well as zoom functionalities for web mapping and mobile applications (adaptive zooming). Update relations allow history management, spatio-temporal analysis and can be a basis for a versioning, which is useful if parallel editing is carried out. If an incremental update is to be carried out automatically, both vertical and update relations are necessary.

Starting from one common relation class the three relation types can be derived and can be specialized further. The unification of all relations has the advantage that common operations on relations can be defined, for example, as "give all features, which are characterized or connected to an element of a relation". The unification further supports the standardized exchange of additional information that goes beyond features and their geometries and attributes. This can be used, for instance, in a Service framework. Finally, a common model of relations supports the persistent storage of the elements of the relations.

There are some significant differences between horizontal relations and vertical respective update relations. First, horizontal relations characterize the states of features, whereas vertical and update relations describe the changes of features. Second, the horizontal relations are modeled as  $n$ -ary relation (order of relationships between  $1 \dots n$ ), whereas the vertical relation is binary (or unary, in case of removal or adding, e.g. matching) and the update relation is also binary (or unary, in case of deletion or insertion).

For the various relation types, approaches for creation are presented and the differences of implementation are discussed. Additionally, methods for visualization are suggested and possible applications are explained. A final example shows the interrelationship during the generalization process which also considers the automated incremental update of the feature sets. Further work, which is currently in progress, will apply the presented concepts on a fully designed map segment.

Future research will investigate relations between relations. For example, one could imagine that a horizontal relation-like an alignment-exists in two different resolutions. Its preservation could be ensured by establishing a vertical relation between these two horizontal relations. In addition, the vertical relation could store relevant meta information on the generalization process. Similarly, it is conceivable to establish an update relation between horizontal relations. Other interesting questions refer to the association of relations and cartographic constraints, for example preserving the density of a partition during generalization. Finally, the model of relations holds potential for the extension or substitution of feature based models.

## Acknowledgement

This research was partly funded by the Swiss Innovation Promotion Agency (KTI/CTI) under the WebGen project (project number 7921.1 ESPP-ES).

## 7 Controlled Line Smoothing by Snakes

Dirk Burghardt

published in  
Geoinformatica Vol. 9, No. 3, 2005, pp. 237-252

**Abstract:** A major focus of research in recent years has been the development of algorithms for automated line smoothing. However, combination of the algorithms with other generalization operators is a challenging problem. In this research a key aim was to extend a snakes optimization approach, allowing displacement of lines, to also be used for line smoothing. Furthermore, automated selection of control parameters is important for fully automated solutions. An existing approach based on line segmentation was used to control the selection of smoothing parameters dependent on object characteristics. Additionally a new typification routine is presented, which uses the same preprocessed analysis for the segmentation of lines to find suitable candidates from curve bends. The typification is realized by deleting undersized bends and emphasizing the remaining curve bends. The main results of this research are two new algorithms for line generalization, where the importance of the line smoothing algorithm lies in the usage of a optimization approach which can also be used for line displacement.

**Keywords.** automated generalization, smoothing, typification, optimization, energy minimization, snakes



## 7.1 Introduction

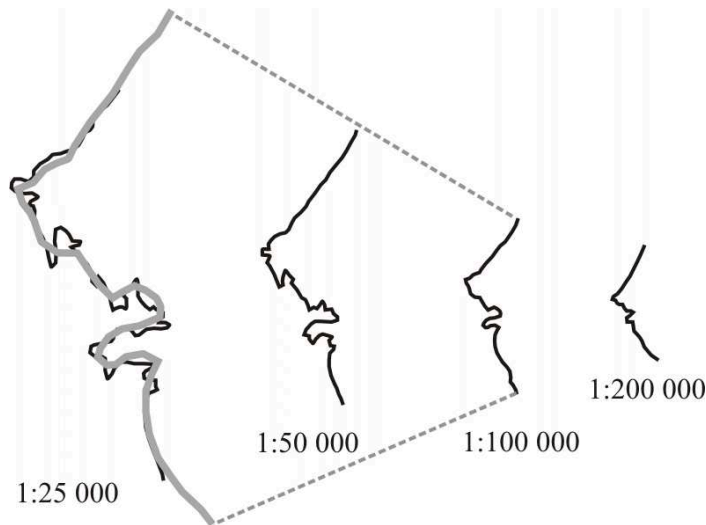
Line generalization applies to different types of cartographic objects, such as angular lines (e.g. canals, buildings), sinuous lines (e.g. hydrography), as well as the depiction of 2.5-D continua (e.g. relief) by contours. Line generalization, just like any other generalization operation, has to observe and preserve the particular characteristics of cartographic objects in the generalization process. Hence, different operators are available for the overall task of line simplification. Angular lines are dealt with by line simplification operators that often rely on a procedure that reduces the number of original points on the line. Line smoothing, on the other hand, is often used for one-dimensional generalization of sinuous lines by 'ironing' away small crenulations (McMaster and Shea, 1992).

The approach presented here focuses on the generalization of sinuous lines. As Figure 1 shows, an approach that was purely based on simplification (i.e., weeding) and smoothing operators, respectively, would not be sufficient, for two major reasons. First, the transitions between angular and sinuous parts of a line are often not distinct. For instance, a border line might contain both angular and sinuous sections, depending on whether it follows survey markers or the centerline of a river. Second, mere point weeding or smoothing simply focuses on single vertices of a line, rather than identifying compound shapes on the line (e.g. bends) and generalizing these shapes (Nakos and Miropoulos, 2003). Plazanet et al. (1998) have presented several algorithms to deal with shape-based generalization of roads. In this article, a method is proposed that is based on an energy minimizing optimization technique, called snakes, that allows to apply controlled line smoothing, while at the same time taking into account the overall shape of the line as well as allowing to integrate different generalization operations.

Because the generalization process follows as a combination of different basic generalization operations, it is always a compromise. In this sense the cartographic solution is generally not unique, but satisfies the different cartographic requirements in a better or worse way. As shown earlier, the use of optimization techniques seems to be suitable for both the combination of different basic generalization operations as well as the control of varying constraints of one generalization operation (Burghardt, 2000).

The use of optimization techniques in the field of automated generalization has been proposed by a number of authors (Bobrich, 1996; Højholt, 1998). Such techniques are primarily applied to the displacement of line objects by means of different approaches – energy minimization methods such as snakes (Burghardt and Meier, 1997; Barrault et al., 2000) and beams (Bader, 2001) and least





**Figure 7.1**  
*Levels of detail for a  
 line object at different  
 scales, after Töpfer  
 (1974)*

squares adjustment (Sester, 2000; Harrie and Sarjakoski, 2000). Sester shows also how to apply optimization techniques for other generalization operations, such as simplification of buildings.

Harrie (2001) extended the least square adjustment approach by inclusion of additional constraint types. The aim of his simultaneous graphic generalization was to have one method for solving different generalization operations in a single optimization step. The new constraint types were concerned with simplification, smoothing and exaggeration. These constraints were based upon pre-computed point movements. For smoothing, he used a Gaussian smoothing approach.

Even though different optimization techniques from several disciplines have been adapted for tasks of cartographic generalization, the resulting linear vector equations after variation or Taylor expansion are all quite similar. Following that, similar methods are used for solving such equations and existing difficulties also are comparable. On the one hand, there is the computational effort required for solving the large equation systems. On the other hand, there are the difficulties to find suitable weights and parameters. Harrie (2001) investigates different strategies (empiricism, machine learning, constraint violation and variance component estimation) with one important restriction: weights have to be determined independently of the shape of objects. This assumption is acceptable only for certain types of constraints, for others, such as curvature constraints, it is important to have different weights for objects of the same type, depending on their shape.

In this paper, a line smoothing approach that is fully integrated with the energy minimization method is first presented. Following that, it is shown how the analysis of shape can help in selecting suitable parameter values for smoothing. An automated controlled line smoothing algorithm is then derived. The same preprocessed analysis is also used to find suitable candidates for typification of curve bends. To improve the recognition of generalized lines a typification procedure with a geometrical basis is suggested.

## 7.2 Energy minimization for smoothing of line objects

The snakes optimization technique used here allows the consideration of different, partly contradictory, generalization constraints. In automated displacement such constraints are the maintenance of minimal distances between objects with correctly represented relative positions and the preservation of typical shapes. Snakes can model these constraints with the help of an energy function consisting of internal and external energies. The internal energy is used to describe the cartographic object's shape and structure. Conflicts, such as distances to other objects that are too small, are calculated by the external energy. In this paper, the snakes model is extended from that originally developed for cartographic object displacement by [Burghardt and Meier \(1997\)](#) by smoothing the shape of line objects, in addition to the internal energy. In this case energy correlates to the degree of detail of the line: the smoother the line the less energy it contains. In combination with automated displacement it would be possible to have one common solution for the generalization of line objects their selection and symbolization. The control of optimization is intuitive, because the external energy refers to other objects, whereas the internal energy describes the line itself.

In this paper the snakes approach will be used in the original form, proposed by [Kass et al. \(1987\)](#). Snakes are energy minimizing splines which adapt their shape and position under the influence of an energy functional. Representing the position of snake parametrically by  $\mathbf{d}(s) = (x(s), y(s))^T$  with arc length  $s$ ,  $s \in [0, l]$ , the energy functional can be written as

$$\begin{aligned} E(\mathbf{d}) &:= \int_l (\mathbf{E}^{int} + \mathbf{E}^{ext}) \cdot ds \\ &= \int_l \left( \frac{1}{2} (\alpha(s) \cdot |\mathbf{d}'(s)|^2 + \beta(s) \cdot |\mathbf{d}''(s)|^2) + \mathbf{E}^{ext} \right) \cdot ds \end{aligned} \quad (7.1)$$

The snake represents a line  $l$  which has to be generalized and the energies are used to model the constraints for generalization. While internal energy  $E^{int}$  has an influence on the shape of the line, the  $E^{ext}$  describes conflicts with objects of the neighborhood. Such conflicts could arise as a result of symbolization, when distances between signatures fall below a minimal distance threshold or overlap each other. There are different requirements for line shape preservation during generalization operations. While displacement should not change too much the shape of the line, smoothing operations may intentionally modify the shape of lines for better visualization quality. With internal energy these different aspects can both be taken into consideration.

For displacement internal energy calculates differences in shape between the original and the displaced line, so minimal internal energy implies minimal deformation of the line. The shape modification is measured by changes of the differences of the first and second derivatives of the original and the displaced line. In case of smoothing internal energy is used to simplify the line. Therefore, first and second derivatives of the line are minimized, resulting in shorter distances between points of vector  $\mathbf{d}$  and minimized curvature of the spline curve. The internal energy terms used for displacement and smoothing, respectively, are formally equal. The differences depend on the definition of vector  $\mathbf{d}$ , which contains the coordinates of the altered line in the case of smoothing (7.2). For displacement (7.3), the differences between the initial and the derived line are used.

$$\mathbf{d}_{smooth}(s) = (x(s), y(s))^T \quad (7.2)$$

$$\mathbf{d}_{displ}(s) = (x(s) - x^0(s), y(s) - y^0(s))^T \quad (7.3)$$

To find the stable state of the snake the functional  $E(\mathbf{d})$  has to be minimized (see Appendix). The variation of  $E(\mathbf{d})$  with constant user-definable parameters  $\alpha$  and  $\beta$  leads to the Eulerian equations, which are solved by discretization in time (Burghardt, 2000; Bader, 2001). The differences in the final formulae between displacement (7.4, 7.5) and smoothing (7.6, 7.7) are straightforward. For smoothing of line objects no external forces are modeled, so the derivatives of external energy  $\mathbf{E}^{ext}$  in  $x$ - and  $y$ -direction are zero.

$$(\mathbf{A} + \lambda \mathbf{I}) \cdot (\mathbf{x}^t - \mathbf{x}^0) = \lambda(\mathbf{x}^{t-1} - \mathbf{x}^0) - \mathbf{E}_x^{ext}(\mathbf{x}^{t-1}, \mathbf{y}^{t-1}) \quad (7.4)$$

$$(\mathbf{A} + \lambda \mathbf{I}) \cdot (\mathbf{y}^t - \mathbf{y}^0) = \lambda(\mathbf{y}^{t-1} - \mathbf{y}^0) - \mathbf{E}_y^{ext}(\mathbf{x}^{t-1}, \mathbf{y}^{t-1}) \quad (7.5)$$

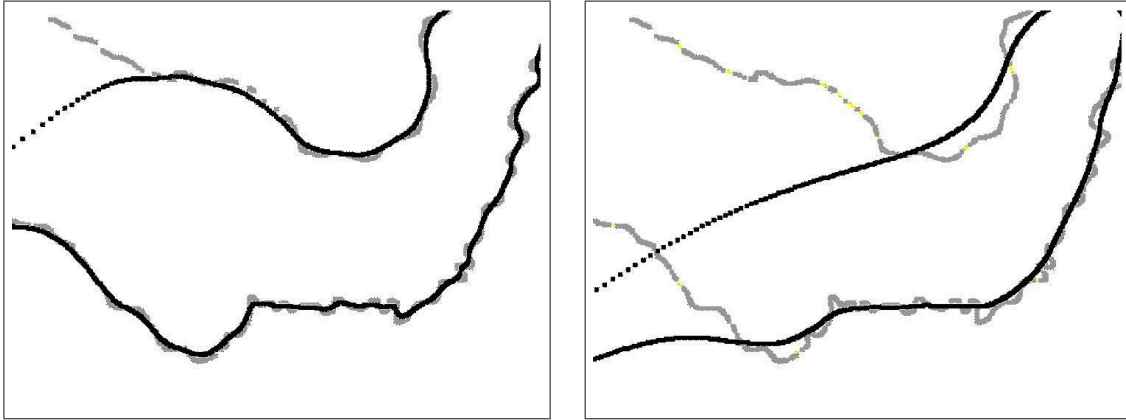
$$(\mathbf{A} + \lambda \mathbf{I}) \cdot \mathbf{x}^t = \lambda \mathbf{x}^{t-1} \quad (7.6)$$

$$(\mathbf{A} + \lambda \mathbf{I}) \cdot \mathbf{y}^t = \lambda \mathbf{y}^{t-1} \quad (7.7)$$

$$\mathbf{A} = \begin{bmatrix} 2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & 0 & \\ -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & \\ \beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta & \dots \\ 0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \\ 0 & 0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta & \\ \vdots & & & & & \end{bmatrix} \quad (7.8)$$

### 7.3 Automated control of parameter selection

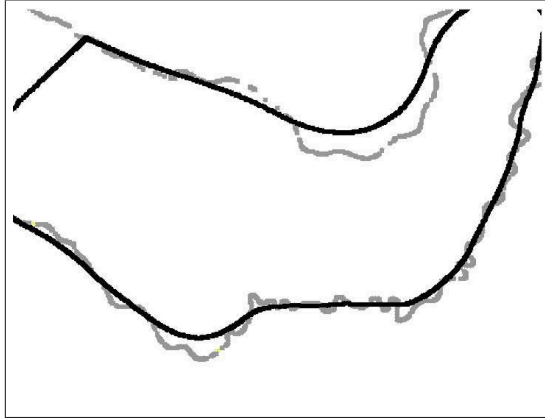
Depending on the degree of smoothing and the distance between points, the end points of the lines are shifted away from their initial position (see Figure 7.2). Higher values of  $\alpha$  produce a stronger minimization of the first term of the internal energy and hence a stronger smoothing and end point shifting (cf. Equation 7.1).



**Figure 7.2**

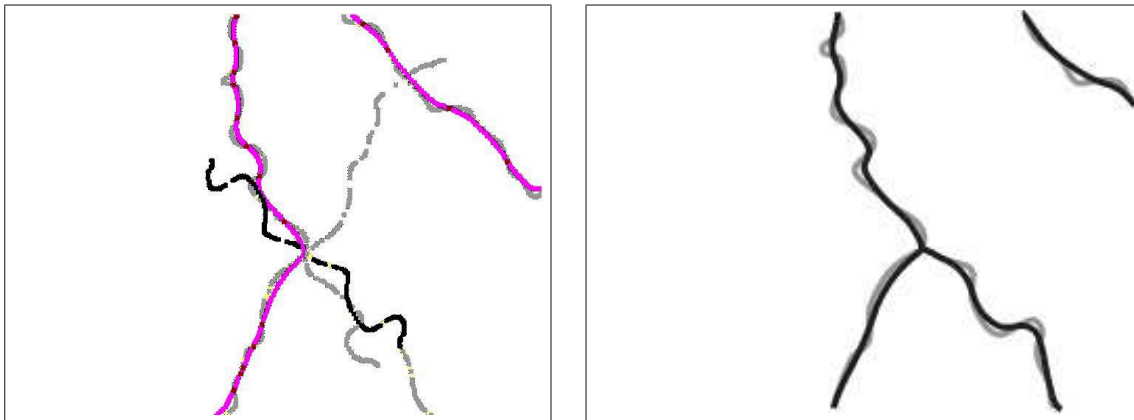
*Excessive shifting of end points of line objects depending on the degree of smoothing.*

To overcome the problem of end point shift there are several counter-measures. One possibility is to include the first and last point of the line multiple times in vector  $\mathbf{d}$ . The points which are added before will be deleted after smoothing. This procedure results in forcing the smoothed line through the end points (see Figure 7.3).



**Figure 7.3**  
*Forced fixed boundary points as a result of introducing multiple end points.*

As a result of internal energy also the curvature at both ends of line are influenced. The line becomes quite straight at their ends because of multiple adding of the same point. That's why a second version of extending the original line was investigated.



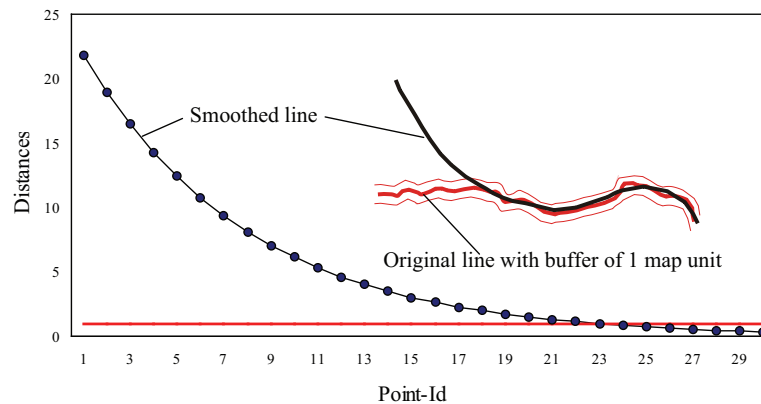
**Figure 7.4**  
*Forced fixed boundary points as a result of using a number of duplicated points at the line end rotated by 180° around the crossroads. Left - Duplicated and rotated points at the line end. Right - Resulting lines after smoothing.*

The idea is to duplicate the segments at either end of the line rotated by  $180^\circ$ . Instead of using the last point multiple times, now some more points at the line end used twice. Thus, the character of the lines is better represented. An important question for any automated solution is the number of additional duplicated points required for bounding the solution. One possible approach which was used here consists in smoothing the line without additional end points and subsequently analyzing the shifted points.

After counting the points for which the distance between the original and smoothed line falls below a threshold value the remaining points up to the end of the line then determine the number of additional end points.

Figure 7.5 shows the distance between the original and the smoothed line for every point on a sample line. In the diagram the distance for the point with Id 24 falls below the threshold value 1.0. That means 23 vertices up to the end of the line were shifted too far. This value will then be increased by 20%, to make sure that the shift affects only the additional points. So, it is recomputed with 23 plus 5 additional end point coordinates. Note that only a fraction of the original line is displayed in the diagram of Figure 7.5; most of the points stay within the threshold distance, as can be seen from the buffer display in the top part of Figure 7.5.

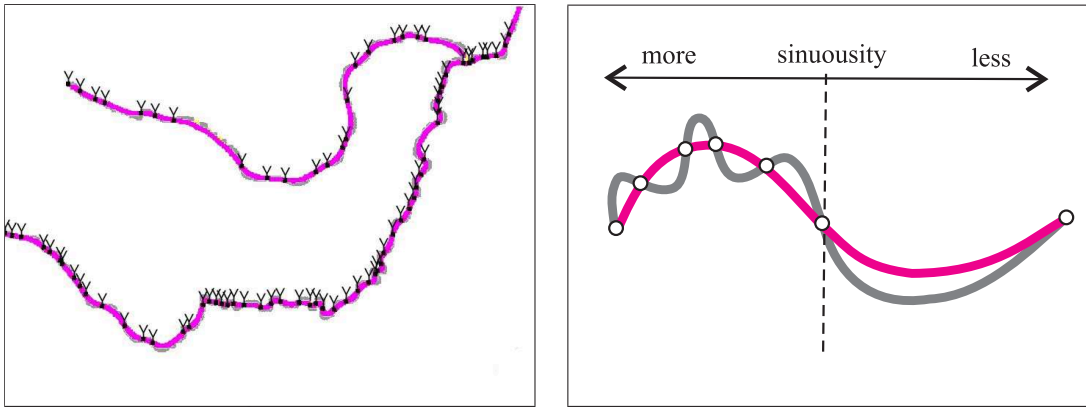
**Figure 7.5**  
Calculation of number of duplicated points depending on distances between original and smoothed line to achieve fixed boundaries.



Before smoothing the line objects a segmentation is necessary to determine the smoothing parameters  $\alpha$  and  $\beta$  of  $E^{int}$ . With the help of the segmentation the lines can be subdivided into smaller segments of different sinuosity. One possibility for segmentation is to smooth the original lines twice, first to determine the characteristics and second with adapted parameter values. The intersection points between the original and the smoothed line correlate with the degree of sinuosity of the original curve.

In Figure 7.6a) the arrows show the intersection points between original and smoothed line. A measure is obtained by counting the intersection points with reference to segment length (Figure 7.6b)). In case the distance between two points of intersection is less than a given threshold value the segment is defined as being sinuous. Note that [Plazanet et al. \(1998\)](#) proposed a similar

approach based on the detection of vertices and the subsequent analysis of the distances between these points to segment the original line. The advantage of the snakes smoothing approach is that it can stay within the same methodological context of energy minimizing optimization techniques. Hence, it does not have to 'piece together' different methodological approaches into one framework, see also [Steiniger and Meier \(2004\)](#).

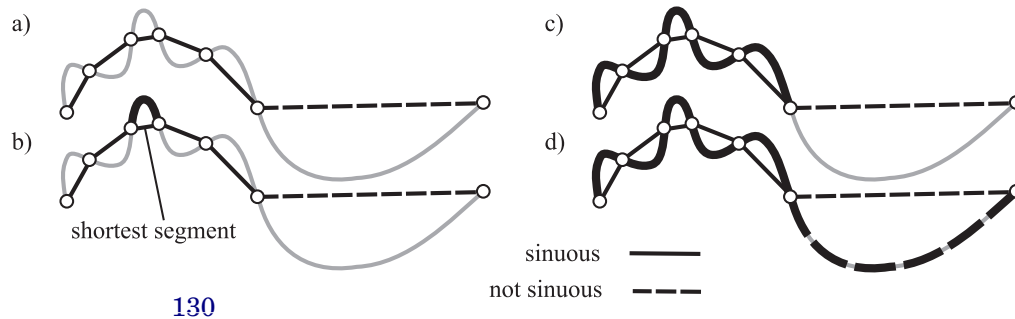


**Figure 7.6**  
Segmentation based on intersection points between original and smoothed line. Left - Arrows show the intersection points. Right - Intersection points as a measure of sinuosity.

Because the sinuosity attribute can change quite often between segments, the next step is to concatenate segments until a user-defined minimal length of line parts is reached. It is necessary to start the procedure with the shortest segments (Figure 7.7b)). If there are adjacent segments with different sinuosity, the longer one determines the value of the sinuosity attribute (Figure 7.7c)). At the last step in Figure 7.7d) the scattered line part keeps the attribute "not sinuous" if the length is longer then the minimal length of line parts defined by the user. In cases where the minimal length is not reached, the scattered line parts take the attribute "sinuous", because the concatenated segments with attribute "sinuous" are longer.

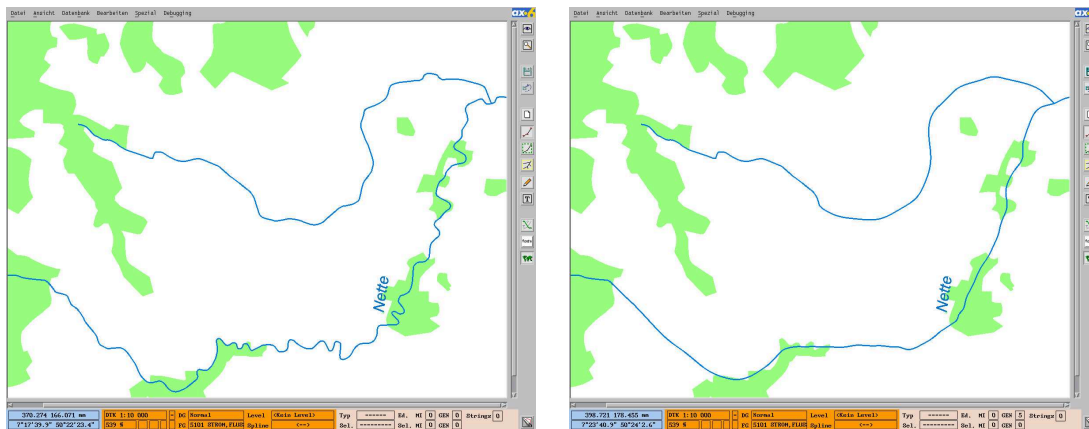
Figure 7.7a) shows the determination of intersection points between the original and the smoothed line. In figure 7.7b) the shortest segment is identified. Figure 7.7c) contains the result of concatenation of segments until a user-defined minimal length of line parts is reached. The final result of a segmented line in sinuous and not sinuous line parts is given in Figure 7.7d).

After segmentation detection the lines are subdivided into the segments and each of them is smoothed with different parameter values for  $\alpha$  and  $\beta$  of the internal energy. Figure 7.7 shows an example in which different parameter values were applied to the seg-



**Figure 7.7**  
Concatenation of  
segments with  
different sinuosity  
attribute.

mented lines. Hence, lines which are sinuous were smoothed more to eliminate the high frequency bends. The less sinuous lines, on the other hand, were less strongly smoothed and the large bends were preserved. However, from a cartographic point of view, this approach can be improved. In order to maintain the more sinuous parts it is necessary to apply typification rules which are described in the next section.



**Figure 7.8**  
An example of the  
use of snakes for line  
smoothing.  
Smoothing was  
applied after line  
segmentation (left -  
original lines, right -  
smoothed lines).

The presented algorithm is integrated in a cartographic production system. The system is used by cartographic experts to create high quality topographic maps. Through the extension of automated parameterization the line smoothing could be carried out also by non-experts, nevertheless the operator would need to decide which classes the smoothing should be applied to and which other generalization operators were needed. The runtime for line



smoothing with snakes is much faster than the line displacement, particularly after subdividing the line in several segments. A disadvantage of the presented approach might be the effort required to implement the matrix equations (7.6,7.7), which can be justified, if a line displacement is also carried out, with the same optimization approach.

The differences between snakes smoothing and other smoothing approaches can be shown on a methodical level (Table 7.1). One classification is made by the filter theory, which uses transformations between spatial and frequency domain. Smoothing after transforming into the frequency domain is realized by frequency filters, e.g. low-pass filters which allows low frequencies to pass. Smoothing within the spatial domain can also be interpreted as filter operations, applied on coordinates instead of frequencies. Detailed analyses of filter characterization for snakes is published by Meier (2000). A second way of classification is by distinguishing between local and global effects of the smoothing algorithms. In the frequency domain algorithms using a Fourier series approach influence the whole line, while Wavelets have a localising component based on their restricted basis functions. Fritsch and Lagrange (1995) have shown that wavelet coefficients are appropriate to characterise the local shape of a curve.

	Spatial domain	Frequency domain
Local	Epsilon filtering Perkal (1966)	Wavelets  Fritsch and Lagrange (1995)  Balboa and Lopez (2000)  Saux (2003)
	Gaussian smoothing Badaud et al. (1986)	
	Sliding average McMaster (1989)	
	Plaster Fritsch (1997)	
Global	Raster  Morphological operations Schweinfurth (1984)	Fourier series  Clarke et al. (1993) Schwarzbach (1995) Fritsch and Lagrange (1995)
	Vector  Snakes (as presented in this section)	

**Table 7.1**  
*Methodical overview of smoothing approaches.*

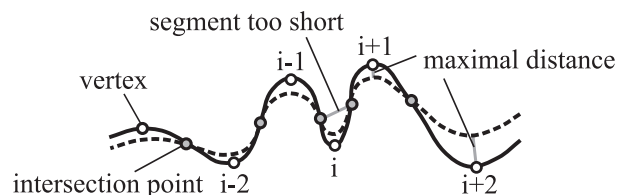
In the spatial domain several smoothing algorithms suggested, which have a local focus. For instance the smoothing approach

suggested by McMaster (1989) considers two or four surrounding points and calculates a straight arithmetic average. In a second step, the actual point is displaced towards the calculated coordinates. Until now smoothing algorithm in the spatial domain with a global focus have only been available for raster data, e.g. the morphological operations (Schweinfurth, 1984). Snakes smoothing fills this gap with their global matrix calculation. The advantage of a global approach is it better preserves the main characteristic of a line, while local adjustment is missing. To obtain a compromise between global and local approaches the automated line segmentation was suggested for controlling the selection of smoothing parameters dependent on object characteristics. While the lines are subdivided into segments of comparable characteristics the approach shifts its focus from a global to a local perspective.

## 7.4 Typification

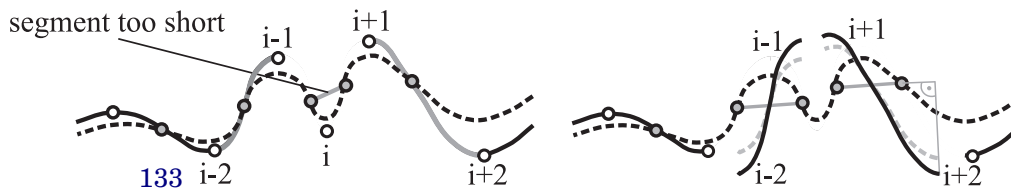
The aim of the typification is to visualize line characteristics even though there are limitations on resolution. For strong sinuous parts this could be achieved by deleting undersized bends and emphasizing the retained or reconstructed ones. To find small bends smoothing is applied the same way as in the first step for segmentation (cf. preceding subsection). Short segments between intersection points of original and smoothed line indicate undersized bends. For typification the intersection points and vertices of the original line have to be calculated. There is one vertex (white) for every bend and every bend is delimited by two intersection points (gray), see Figure 7.9. The vertex of one bend is the point with the maximal distance between the original (black) and the smoothed line (dashed).

**Figure 7.9**  
Calculation of vertices  
(white) and  
intersection points  
(gray).



Starting with the vertex (i) of the undersized bend, all line vertices between the previous (i-1) and following vertex (i+1) of the original line are deleted (Figure 7.10a). The vertices of the adjacent bend sides, between vertex i-1 and i-2 as well as i+1 and i+2, respectively, are moved for the construction of the new, typified bend. Its direction and length are calculated from the connection

of intersection points between the original and the smoothed line. The translation leads in the direction of the deleted bend with a value of half the intersection point distance. The dotted gray lines in Figure 7.10b show the intermediate step.



**Figure 7.10**  
*Construction of new curve bends through translation and exaggeration.*

An exaggeration of the newly constructed bends is also possible with a distance dependent stretching of line vertices perpendicular to the connecting segment between intersection points. The solid black lines in Figure 7.10b show the exaggerated bends. Finally, the original line with its constructed new bends is smoothed by energy minimization. The second term of the internal energy guarantees the continuity of the smoothed line. Also, the distance between vertices of the line becomes approximately equidistant as a result of the first term of internal energy.

Figure 7.11 shows two examples of typification of line objects from VECTOR25 road network. The road network of VECTOR25 was digitized on the basis of topographic maps with scale 1 : 25'000. The aim for this example was to generalize the road network for a smaller scale 1 : 500'000. In the background of road objects the corresponding manual generalized topographic map with scale 1 : 500'000 is shown. On the left side you can see the situation before on the right side after typification.

The circles bring out the results of typification of object class "Main road", which are in the second example very similar to the manual solution. More examples are shown in Figure 7.12 for another object class "Road open to traffic" of road network. It can be seen that not all sinuous line segments would be typified (double encircled curve sequence). The main reason is that calculation of vertices and intersection points is dependent on preprocessed smoothing. In this example the preprocessed smoothing was stronger, so in the double encircled area no intersection points are calculated between original and smoothed line.

To overcome this problem a strategy could be used which implies a frequency dependent typification. If preprocessed smoothing is not so strong, only smaller bend sequences are typified (Figure 7.4b)), if a stronger preprocessed smoothing is used the longer bends (relating to long wavelength) will be typified (Figure 7.4c)).

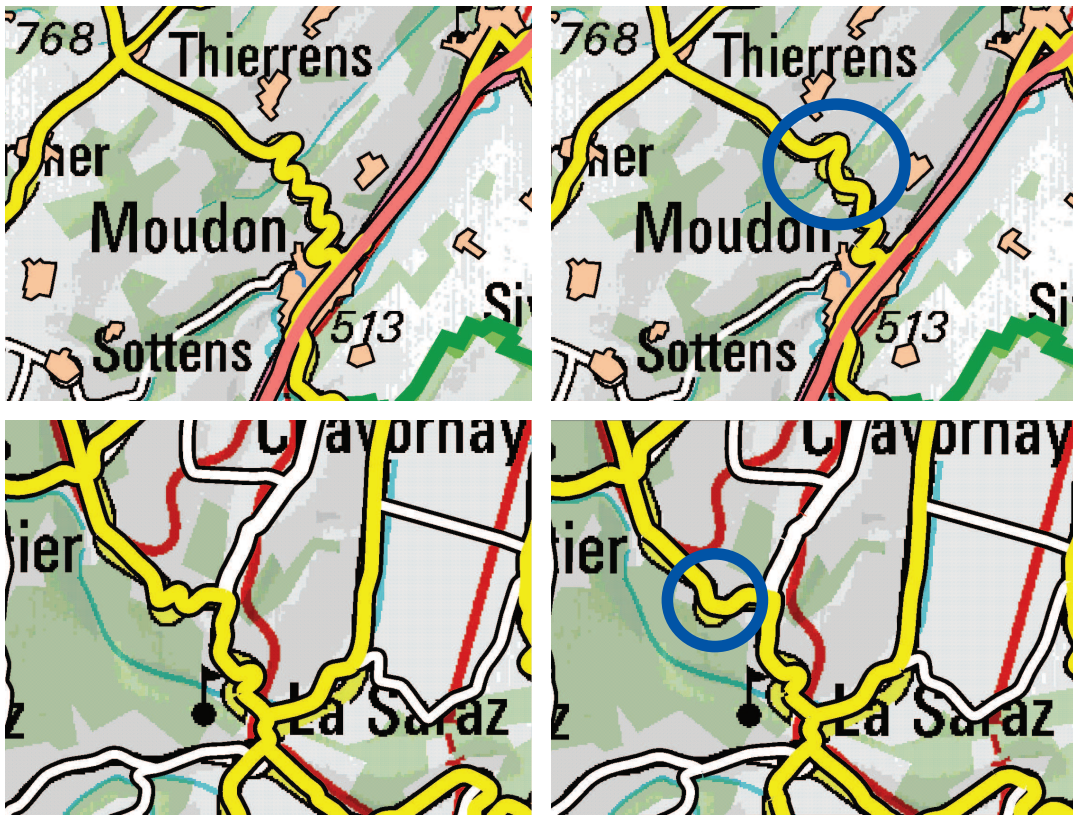


Figure 7.11

Typification  
of road

topogr

backgr

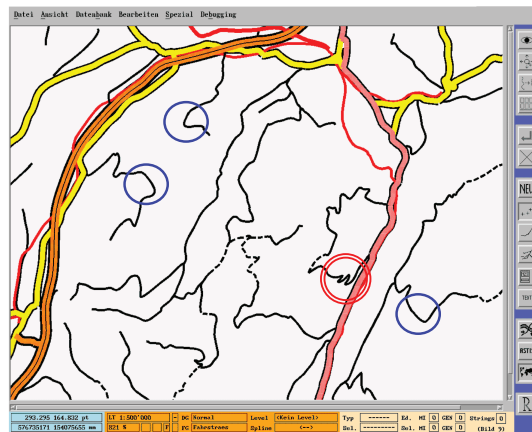
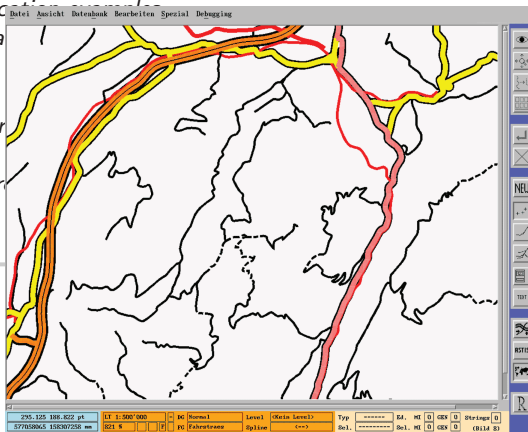
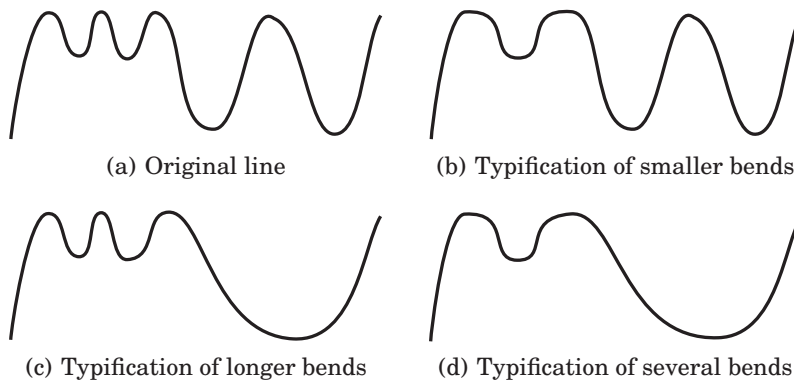


Figure 7.12

Typification and  
smoothing of road  
network.



**Figure 7.13**  
Frequency dependent  
typification with  
different amounts of  
preprocessed  
smoothing.

In general an iterative typification with different amounts of preprocessed smoothing could be applied.

An alternative to this typification approach is the Accordion algorithm suggested by [Plazanet \(1996\)](#). This algorithm aims to enlarge a bend or bend series to remove the bends that coalesce. The central inflexion point of the line has to be fixed and all the others points are moved away from it, specific to every bend, in the orthogonal direction of each bend axis. The main difficulty of this approach is to avoid creating new conflicts when solving the initial ones. Further research ([Ruas, 2000](#); [Duchêne, 2001](#)) has been introduced micro and meso Agents to overcome this side effects.

Compared with the Accordion algorithm, the typification presented here approach has the advantage that no side effects were produced, because bends will be removed instead of displaced. A negative consequence is that less information will be visualized in the map. In this sense the typification approach suggested here can be used as a complementary procedure to the Accordion algorithm. If the local situation allows visualizing of all curve bends, the Accordion algorithm makes sure that the bends do not overlap and if too much side effect occurs then our typification approach can be used to reduce the number of curve bends.

## 7.5 Conclusion

The results of the research presented here are two new algorithms for line smoothing and typification. Advantages of line smoothing with snakes are discussed from a methodical and a practical point of view. Snakes smoothing fills the gap of smoothing operators with global focus in the spatial domain. The advantage of a global approach is the better preservation of the main charac-



teristic of a line. From a practical point of view the snakes approach uses an optimization approach which can also be used for line displacement. The advantage of the snakes model is the simple combination of several constraints with the help of different energies. External energy describes conflicts with other map objects, while internal energy models the shape constraints of cartographic lines. The main difference of using snakes for displacement or smoothing, respectively, depends on the definition of the internal energy. While the shapes of lines should be preserved during displacement, smoothing implies more considerable deformations.

To improve the results of smoothing, the lines can be smoothed twice. A first pass is executed with default parameters for  $\alpha$  and  $\beta$ , independently of line characteristics. The resulting line crosses the original line and the density of intersection points between the two lines provides a measure of sinuosity. A subsequent segmentation helps to determine parts of the line with similar characteristics. For the second smoothing pass the parameters  $\alpha$  and  $\beta$  can then be selected in relation to the line shape established previously. For instance, lines which are more sinuous can be smoothed more strongly to eliminate high frequency bends. The reason for using one common approach for smoothing and displacement is the easier control of the interaction between displacement and smoothing. Additionally, it is faster to apply smoothing and displacement together than one after another. Further work should investigate smoothing with position dependent parameter  $\alpha = \alpha(s)$  and  $\beta = \beta(s)$ , to achieve a more local control of smoothing.

Additionally, a typification routine with a geometrical basis is suggested. Comparing with other algorithm the approach has the advantage that no side effects were produced. It is well suited for greater scale transitions and can be used as supplementation for Accordion algorithm. In general the algorithm works well, but depending on the degree of the preprocessed smoothing different curve bends for typification are selected. The consequence is that not all curve bends are typified in one step and an iterative process has to be applied, which leads to a frequency dependent typification.

## 8 Mesh Simplification for Building Typification

Dirk Burghardt and Alessandro Cecconi

published in  
International Journal of Geographical Information Science,  
Volume 21, No. 3-4, 2007, pp. 283-298

**Abstract.** This paper describes an approach for the typification of buildings using a mesh simplification technique. The approach is adapted from the area of computer graphics and was originally developed for surface reconstruction and mesh simplification. The main goal was to develop an algorithm which creates fast and reproducible results. The typification procedure is modeled as a two-stage process, with the steps 'positioning' and 'representation'. While the positioning step determines the number and the position of the building objects based on Delaunay triangulation, the representation step is used to calculate size and orientation for the replacement buildings. The presented results show the important influence of weights during positioning steps to control the object distribution. Proposed parameters are the number of objects as well as several object characteristics such as size, shape, orientation and semantic. The approach has to be extended, if also building alignments should be preserved. Further applications are imaginable, for instance the icon placement on dynamic maps.

## 8.1 Introduction

The operator typification is used for the transformation of an initial set of objects into a subset, while maintaining the distribution characteristics and pattern of the original set. A first step can be an identification of clusters (Regnauld, 1996; Anders and Sester, 2000; Anders, 2003), which then allows the preservation of important cartographic structures like alignments of buildings along roads during the reduction of object numbers. Several clustering algorithms are proposed especially for data analysis and information retrieval (Guha et al., 1999; Karypis et al., 1999; de Berg et al., 2004) also with support of Delaunay data structures to incorporate spatial proximity (Kang et al., 1997; Estivill-Castro and Lee, 2000; Lee and Estivill-Castro, 2002), but there are less applications to generalization with consideration of cartographic constraints (Jones et al., 1995; Jones and Ware, 1998; Mustière and Moulin, 2002; Burghardt et al., 2004b).

Kohonen Feature Maps, a neural network learning technique, used especially for the typification of 2D-structures of similar type and size like buildings (Højholt, 1995; Sester and Brenner, 2000; Allouche and Moulin, 2002). The prominent property of this unsupervised learning method is the fact that the neurons are adapted to a new situation, while keeping their spatial ordering. The approach is non deterministic as a result of random selection of neurons at the beginning, which means after rerun the algorithm different results will be achieved. We propose a deterministic approach for the typification of the feature class building with adaptation of a method from computer graphics, so-called mesh optimization.

## 8.2 Theory of mesh optimization

The mesh simplification technique is related to mesh optimization techniques well known in the research area of computer vision which hold great potential for many applications. The main reason for using a mesh optimization technique is to reduce the amount of data. Thereby three goals or purposes can be defined:

- i) faster rendering
- ii) reduced storage volume
- iii) simpler manipulation

Mesh optimization, as considered by Turk (1992) and Schroeder et al. (1992), refers to the problem of reducing the number of faces in dense meshes (usually made up of triangles). The research work



of Hoppe et al. (1993) has shown that mesh optimization can be put into use effectively in at least two applications: surface reconstruction from unorganized points, and mesh simplification (the reduction of the number of vertices in an initially dense mesh of triangles).

Their principal idea is to describe mesh simplification as an optimization problem, defining an energy function  $E$  that directly measures deviation of the final mesh from the original. To solve the optimization problem they minimize the energy function  $E$  that captures the competing objectives of a tight geometric fit and a compact representation. One of the main disadvantages concerning this approach is the time-consuming computation of the energy minimizing function  $E$ .



**Figure 8.1**  
*Example of reducing the number of vertices of a mesh (Hoppe et al., 1993).*

Fig. 8.1 illustrates the main idea of reducing the amount of vertices of a mesh (mesh simplification) for multiple representations preserving the original characteristic vertex distribution. On the left hand side of Fig. 8.1, a dense mesh made up of a large number of vertices describes the front of a 'face'. In the middle part, the amount of vertices is reduced maintaining the typical outline and characteristic of the original form. On the right hand side a strongly simplified version, with a fraction of the original vertices of the starting data set, is depicted. Each state best represents the original form. Thus the selection of points plays a crucial role and must be optimized to preserve the original arrangement. Regions with a dense number of vertices should be maintained in each state or level as dense areas and vice versa with sparse parts. This is well illustrated in Fig. 8.1. Transferring this idea to cartography and particularly for building representation, the various states of

the 'face' in Fig. 8.1 can be looked upon as different scales or LODs of a map. The vertices do thereby not display the buildings themselves but, for example, the centers of gravity of building objects.

Starting from a mesh  $M_0$  with  $n$  vertices (representing the centers of gravity of the individual buildings) a new mesh  $M_j$  with  $(n - 1)$  vertices is sought. This new mesh should best represent the original one with minor changes. Hoppe et al. (1993) define a mesh  $M$  as a pair  $(K, V)$ , where:  $K$  is a simplicial complex representing the connectivity of the vertices, edges and faces;  $V = \{v_1, \dots, v_m\}, v_i \in R^3$  is a set of vertex positions defining the shape of the mesh. To obtain a mesh that provides a good fit to the original point set  $X$  an energy function  $E(K, V)$  is defined where:

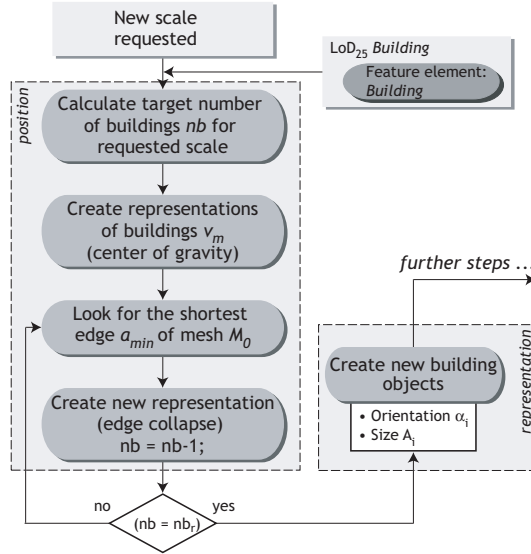
$$E(K, V) = E_{dist}(K, V) + E_{rep}(K) + E_{spring}(K, V)$$

By varying number, position and connectivity of the vertices a minimization of this value is looked for. The distance  $E_{dist}$  is equal to the sum of squared distances from the point set  $X$  of the mesh. The value  $E_{rep}$  is proportional to the number of vertices.  $E_{spring}$  is looked upon as a regularizing term and describes the sum of the edge lengths. Using this energy function  $E$  for a mesh optimization in principle provides the possibility of observing several constraints which restrict the generalization process (such as the distance between the objects, building alignment, etc.). But since, on the one hand, the minimization of the energy function  $E(K, V)$  is computationally very time-consuming (Puppo and Scopigno, 1997) and on the other, only *one* constraint (distance between the objects) has to be considered in this work, a simpler function, more adapted for on-demand web mapping has been derived.

### 8.3 Mesh simplification adapted for building typification

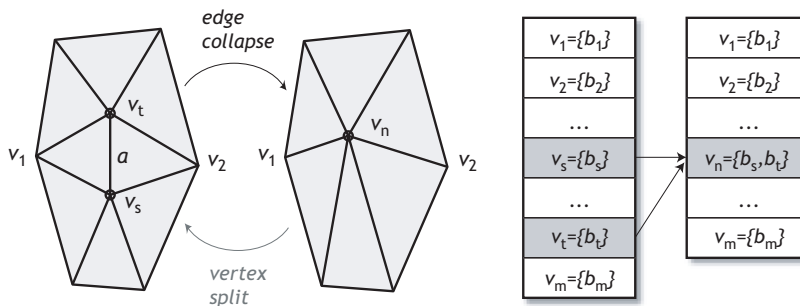
Before explaining the various steps of the proposed mesh simplification technique in more detail, the full approach of typification should be discussed. The typification procedure, as shown in Fig. 8.2, is composed of two steps which are not independent and interact with each other:

- **Position:** Determining the number and the placement of the new objects with respect to the requested scale;
- **Representation:** Creation of a new building objects at the determined position.



**Figure 8.2**  
Flowchart for the typification operator of the generalization process for the feature class building, carried out by the mesh simplification technique. The highlighted part on the left computes the position of the new placeholders, while the right part defines the representation.

The typification process (mesh simplification) is an iterative process where the termination criterion is dependent on the requested map scale  $m_r$ . By means of the original number of objects  $nb_o$  and the scale value  $m_r$ , an adapted number of building objects  $nb_r(nb_o, m_r)$  is computed for terminating the iteration. The basic idea is, that two objects which lie next to each other can be replaced by a new one between them, called a representative (placeholder). This simple operation is known as 'edge collapsing' (Hoppe, 1996) and is sufficient for effectively simplifying meshes. As shown in Fig. 8.3, an edge collapse transformation unifies two adjacent vertices  $v_s$  and  $v_t$  into a new single vertex  $v_n$ .



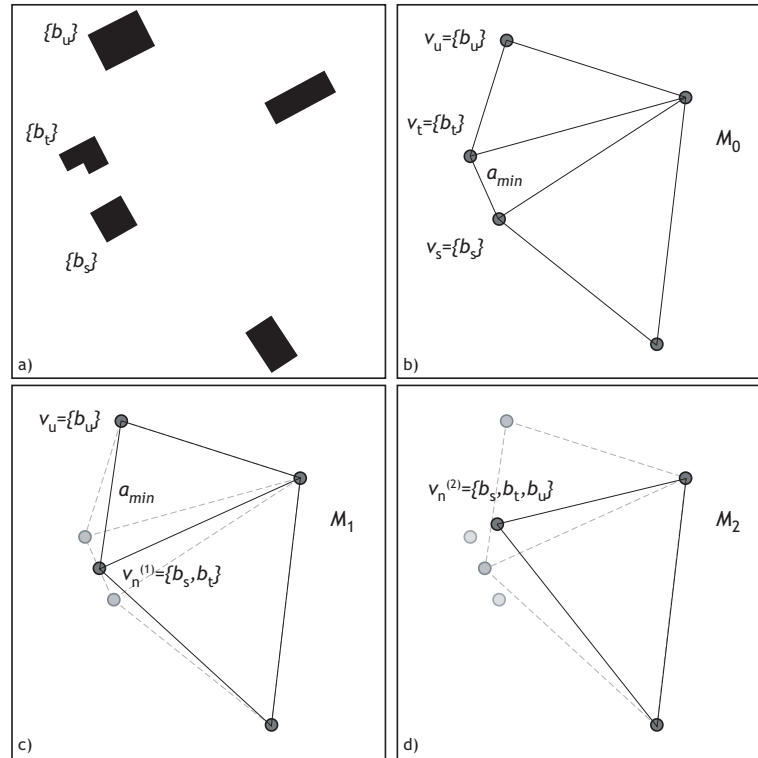
**Figure 8.3**  
The method of edge collapsing for mesh simplification (left). On the right a sequence of edge collapses is shown.  $v_i$  describes the center of gravity of building  $b_i$ .

The inverse transformation vertex split adds a new vertex  $v_t$  and thus two new faces to the original mesh. Since typification must reduce the number of objects, the inverse case is not relevant here. Transferring the idea of edge collapsing to the feature class building, where each vertex represents a building object, helps to solve the first step of the typification process - 'position'. For the second step 'representation', each remaining vertex  $v$  of the final mesh  $M_f$  must know which building or buildings it represents. For example, in Fig. 8.3 the placeholder  $v_n$  represents the vertices  $v_s$  and  $v_t$  and thus the building objects  $b_s$  and  $b_t$ . From the geometric information of the represented objects (e.g. area  $A$ , orientation  $\alpha$ ) a new best fit building object (i.e. a placeholder or representative) must be created for the requested map scale  $m_r$ .

### 8.3.1 Positioning

The main phases of the mesh simplification technique for the feature class building are illustrated in Fig. 8.4. It explains the first step of the typification process (position).

**Figure 8.4**  
Phases of the mesh  
simplification  
technique for the step  
'position'.



Let  $X$  be a set of points with vertices  $v_1, \dots, v_m$  representing the centers of gravity of the building objects  $b_1, \dots, b_m$  in  $\text{LOD}_{25}$  and  $M_0$  a mesh over this point set. In a first iteration loop a new mesh  $M_1$  is sought, where two vertices  $v_s$  and  $v_t$  are replaced by a new one  $v_n^{(1)}$  as illustrated in Fig. 8.4c). The criterion for the selection of these two vertices is that they describe the shortest edge  $a_{\min}$  of the mesh  $M_0$ . The new vertex  $v_n^{(1)}$  is defined as center of  $a_{\min}$  and thus lies between  $v_s$  and  $v_t$ . By searching the shortest edge and replacing  $v_s$  and  $v_t$  through  $v_n^{(1)}$  the modifications take place locally and thus the main characteristics of the mesh will not be disturbed.

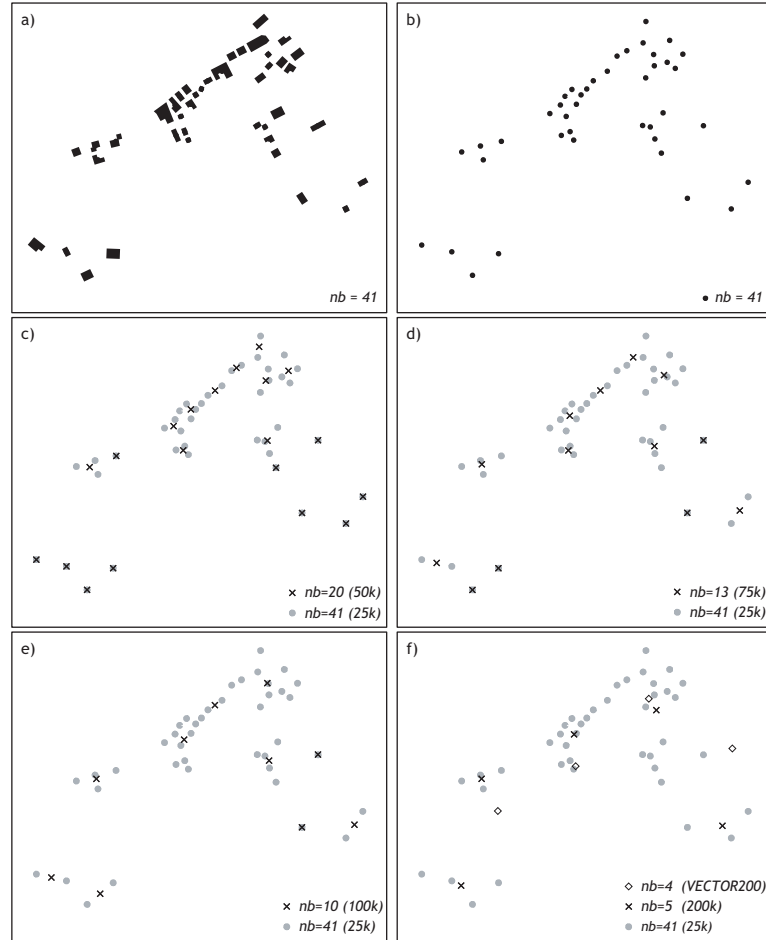
For the calculation of the position of the new vertex  $v_n^{(1)}$  all affected vertices (and thus buildings) must be considered:  $v_s$ ,  $v_t$  and  $v_u$ . Hence, the position of  $v_n^{(2)}$  is fixed by the center of gravity of the vertices which are replaced (shown in Fig. 8.4d)). The iteration is continued until the current number of vertices meets a threshold value  $nb_r$ , which is dependent from target scale.

After termination this iteration process each remaining vertex represents one or more of the original buildings. In the following figures the results computed with the mesh simplification technique are illustrated. A comparison of original objects and created placeholders illustrates the obtained results.

Fig. 8.5a) shows the building objects of  $\text{LOD}_{25}$  (VECTOR25) with  $nb = 41$  elements. In Fig. 8.5b) only the centers of gravity representing the original data set are displayed. These points compose the vertices of the mesh  $M_0$  as discussed in the previous section and define the starting point of the mesh simplification technique. With the iterative process of *edge collapsing* a number of vertices are removed or replaced by new ones. Fig. 8.5c) shows the result for the requested scale of 1 : 50'000 displayed as black points whereby the amount is decreased to  $nb = 20$ . In the background the vertices depicted in gray describe the vertices of  $M_0$ . In some cases vertices coincide and only the black ones (1 : 50'000) are visible. The placeholders for scale 1 : 75'000 are displayed in figure 5d) where the number decreases to  $nb = 13$ . In Fig. 8.5e) the mesh is composed of only  $nb = 10$  vertices. Finally, in Fig. 8.5f) a comparison with the positions of the vertices of  $\text{LOD}_{200}$  can be done. In contrast to  $\text{LOD}_{200}$  the computed solution results in 5 objects, whereby three of them are describing a similar position. Evaluating the created placeholders for each scale with the original data set of  $\text{LOD}_{25}$  it can be noticed that reasonable results have been achieved maintaining the main characteristics of the group through all scales.

**Figure 8.5**

The positions of the placeholders for the following scales: a) buildings in  $LOD_{25}$ , b) 1:25'000 (original data set), c) 1:50'000, d) 1:75'000, e) 1:100'000 and f) 1:200'000 with the objects' center of gravity in  $LOD_{200}$  (all not to scale). Data: VECTOR25/200 © swisstopo (BA034957).



### Stopping criteria for removal of objects

The most important constraint for the removal process is to keep the 'black-white' ratio constant. If the positioning and representation phases are combined in one iteration step, the ratio can be easily computed. Assuming that the object size in the target scale is bigger, because objects have to be enhanced to maintain legibility, the removal stops if the sum of area at source and target scale is the same. The calculation of object size will be shown later, when explaining the representation phase.

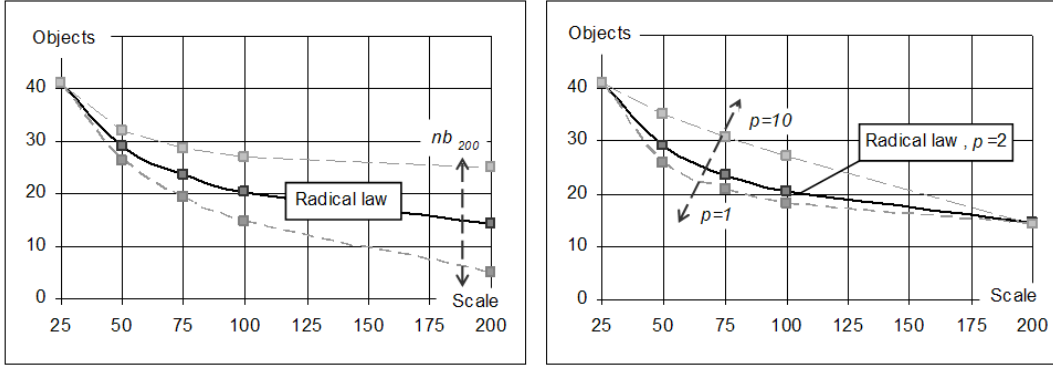
To determine the number of objects for the target scale independent from representation phase Töpfer's radical law (Töpfer and Pillewitzer, 1966) can be applied:

$$nb_r = nb_{25} \sqrt{\frac{m_{25}}{m_r}}$$

Comparing the results with manually generalized VECTOR25/-200©swisstopo data the results did not match very well. For example at scale 1:200 ( $m_{200}$ ) Töpfer's radical law forecasts three times more objects than were found in manually generalized VECTOR200 data. Therefore an interpolation function based on an extension of Töpfer's radical law was applied to calculate the number of objects  $nb_r$  for the intermediate scales ( $m_r$ ) from two source scales, in our case 1:25k ( $m_{25}$ ) and 1:200k ( $m_{200}$ ):

$$nb_r = nb_{25} \left( \frac{m_{25}}{m_r} \right)^{\frac{1}{p}} \left( 1 - \frac{m_r - m_{25}}{m_{200} - m_{25}} \right) + nb_{200} \left( \frac{m_{200}}{m_r} \right)^{\frac{1}{p}} \left( \frac{m_r - m_{25}}{m_{200} - m_{25}} \right)$$

While Töpfer's radical law uses the number of objects at one source scale ( $nb_{25}$ ) to calculate the number of objects at the target scale ( $nb_r$ ), our extension allows an interpolation between two source scales as given in a Multi-Representation Databases (MRDB). The additional parameter is the number of objects ( $nb_{200}$ ) at the second scale (see figure 6 left) in our case given through VECTOR200 data.



With the power  $1/p$  a second extension of Töpfer's radical law was applied to modify the increase of the interpolation curve. The power  $1/p$  has to be greater than one. For  $p = 2$  the radical law is obtained (see Fig. 8.6 right).

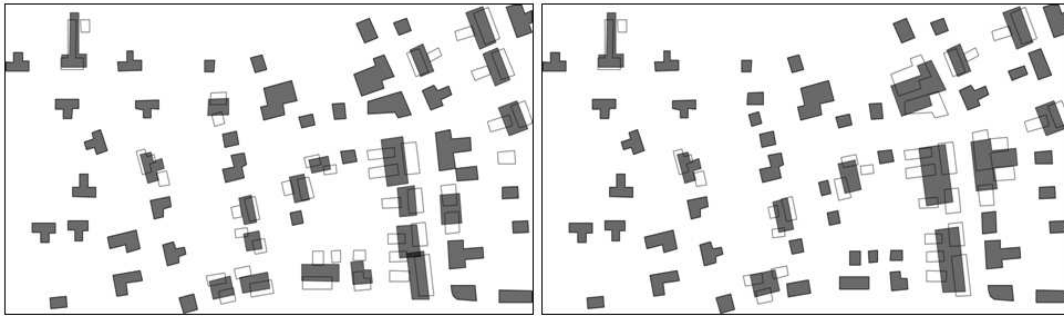
#### Distribution of objects

After determination of how many objects should be presented, the other important questions are which objects should be removed

**Figure 8.6**  
Interpolation function based on an extension of Töpfer's radical law with two parameters - number of objects ( $nb_{200}$ ) at a given second scale (left) and the power  $1/p$  to modify the increase of the curve (right).



and where to place their representatives. The basis for these decisions are the lengths of the Delaunay edge. For performance reasons the Delaunay triangulation is calculated from the centroids of the objects.



**Figure 8.7**

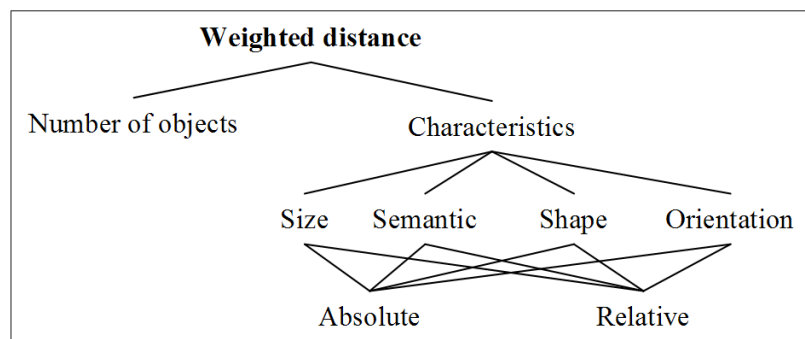
*Mesh simplification based on Delaunay triangulation removing shortest edges between centroids (left) or edges with shortest distance between the objects outline (right).*

Following that the Delaunay edges are ordered for a collapse operation based on length between centroids or shortest distance between the objects belonging to it. The differences in the results are shown in Fig. 8.7. Using shortest distances between objects creates more homogeneous results.

Calculation time for Delaunay triangulation and the ordering of the Delaunay edges are the limiting factors of the algorithm. Tests were carried out with a normal PC (Intel Pentium processor 1.4 GHz), whereby the triangulation took 100 ms for 104 objects and 5'147 ms for 1'067 objects. The used sort-function for the Delaunay edges has a runtime of  $n \log(n)$  and needs 10 ms (291 edges) respective 150 ms (3161 edges) in the same test scenario.

**Figure 8.8**

*Weighting of Delaunay edges to influence distribution of objects.*



The length of Delaunay edges can be weighted in different ways as shown in Fig. 8.8. An important factor is the number



of objects that a vertex represents. The second main category are the objects characteristics such as size, semantic, shape or orientation, which can also be considered for weighting of object removal and placement of representatives.

**Weighting through the number of objects** Since our approach is based on looking for the vertices describing the shortest distance (shortest triangle edge), areas with high building density would be thinned out more strongly than sparsely populated regions. On the one hand this could be desirable to avoid too many objects in urban areas, but on the other hand it would distort the characteristics of an area. To meet this requirement a local density factor  $f_a$  is included in the calculation of the distance. This factor  $f_a$  allows to elongate the real edge length  $a$  between the vertices and thus to decrease the thinning process in densely populated areas. It depends on the number of objects  $r$  that a vertex represents (for example the value for  $v_n^{(2)}$  is  $r = 3$ ) as well as the term  $s_u$  which must be set by the user in advance meeting the following conditions:

$$0 \leq s_u \leq 1 \quad \longrightarrow \quad \begin{array}{ll} s_u = 0.0 : & \text{no correction is done} \\ s_u = 1.0 : & \text{correction with the number} \\ & \text{of represented objects} \end{array}$$

From these two values the factor  $f_a$  can be defined as follows:

$$f_a = s_u \cdot (r - 1) + 1.$$

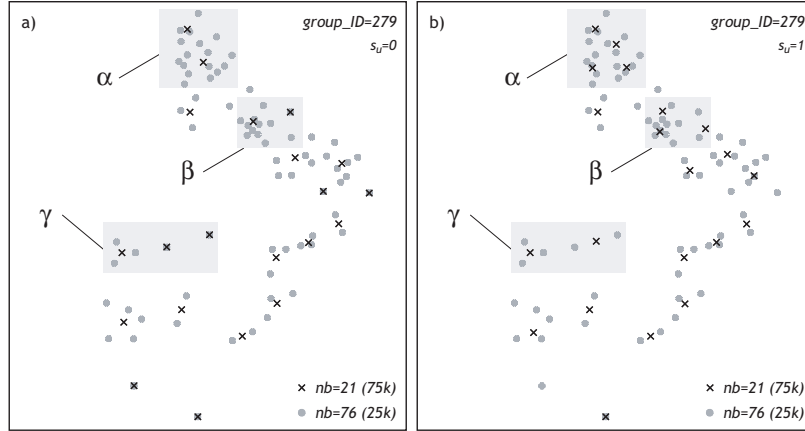
As described above the shortest edge  $a_{min}$  of all edges  $a_i$  ( $\forall i$ ) of the mesh  $M_0$  is looked for:

$$a_{min} = \min[a], \quad \text{where} \quad a = f_a \cdot \sqrt{(x_s - x_t)^2 + (y_s - y_t)^2}.$$

The examples discussed so far have been computed with the factor  $s_u = 0.0$  which implies that no correction has been done concerning the distance between the vertices. To incorporate this quality the value  $s_u$  has been introduced to decrease the thinning process in densely populated areas. The difference between computing a representation with  $s_u = 0.0$  or  $s_u = 1.0$ , respectively is displayed in Fig. 8.9a) and b) for the scale 1:75'000 ( $s_u$  may range between 0.0 (no correction) and 1.0 (maximum correction)).

For both results the number of objects to be displayed for the requested scale is equal ( $nb = 21$ ). What changes are the positions of the centers of gravity of the placeholders. On the left side of Fig. 8.9a) the highlighted area  $\alpha$  is represented by only two buildings while on the right part of the Fig. 8.9b) four objects are displayed

**Figure 8.9**  
The positions of the placeholders with  $s_u = 0.0$  (left) and  $s_u = 1.0$  (right), both not to scale. Data: VECTOR25 © swisstopo (BA034957).



within the same area. Comparing these two areas reflects the effect of the factor  $s_u$ : computing the result with a high value for  $s_u$  ( $s_u \rightarrow 1.0$ ) takes the density of the building objects more into account than with low values ( $s_u \rightarrow 0.0$ ). The same effect can be found in the second highlighted example  $\beta$  where places with a high density of buildings are represented better in b) than in a). Working with low values for  $s_u$  ( $s_u \rightarrow 0.0$ ) prefer single seated objects as illustrated in the highlighted area  $\gamma$ .

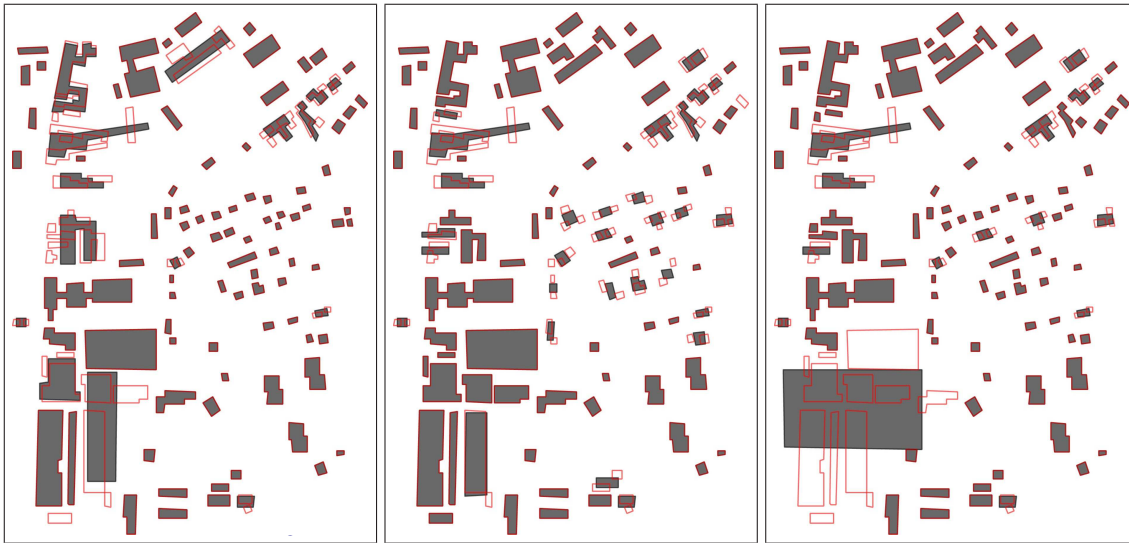
**Weighting through consideration of object characteristics** An alternative weighting also allows the consideration of object characteristics such as size, shape, orientation and semantics. For all properties relative or absolute values can be calculated (see Table 8.1). While relative weighting means only the characteristics of the objects belonging to the Delaunay edge are compared, absolute weighting helps to preserve certain objects characteristics in general.

**Table 8.1**  
Weighting through consideration of object characteristics.

	Absolute	Relative
<b>Size</b>	weighting with sum of object areas ( $A$ ) means smaller objects will be replaced first ( $w = A_1 + A_2$ )	weighting with quotient of areas means objects of similar size will be aggregated first ( $w = A_2/A_1$ if $A_1 > A_2$ or $w = A_1/A_2$ if $A_2 > A_1$ )
<b>Shape complexity</b>	preferred replacement of objects with simple shapes	preferred replacement of objects with similar shapes
<b>Orientation</b>	replacement of objects, which have diagonal orientation	replacement of objects with similar orientation
<b>Semantics</b>	objects of a certain class will never be deleted	replaced objects must belong to the same class

For instance a weighting through the consideration of shape complexity can be applied to keep objects with a conspicuous and more complex shape. Measures to calculate the shape complexity

are described in AGENT (1999). The semantics of the objects restricts also the typification process either relatively or absolutely. Relative restrictions are for example that the replaced objects must belong to the same class, absolute restrictions are that, for instance, objects of a certain class will never be deleted (e.g. public buildings such as schools or hospitals).



**Figure 8.10**  
Comparison of unweighted object selection (left) with weighted selection considering absolute size (middle) and size differences (right).

A weighting with the size of the object is shown in Fig. 8.10. The case of considering the size through a weighing proportional to the area of the involved objects leads to a replacement of smaller objects and prevents an replacement of bigger objects, which is our preferred solution (absolute case). Alternatively the size differences of objects belonging to the Delaunay edge can be used, whereby objects with large size differences will not be represented through one representative. The disadvantage is seen in Fig. 8.10 right, where for instances some bigger objects are typified through one representative, while a small object near by is preserved.

#### Representation - shape construction of replacement buildings

The second step of the typification process named *representation* creates for each vertex a building object from the ones it represents. For example in Fig. 8.4, for the vertex the vertices  $v_s$ ,  $v_t$  and  $v_u$  determine the shape of the new building object  $b_n^{(2)}$ . A simple approach calculates the area from the old buildings preserving the "black-white-ratio" and uses the *shape* and the *orientation* from

the biggest ones. Results are shown in Fig. 8.7 and Fig. 8.10. An alternative approach attempts to maintain the characteristics of the whole group. Thereby the two attributes *area*  $A_{[width, length]}$  and *orientation*  $\alpha$  of each represented object are considered.

**AreaA** Comparing the map series of the Swiss National Mapping Agency it can be stated that most buildings are strongly simplified at a scale of 1:100'000 and smaller and thus less detailed. The shapes are usually represented as rectangles to meet the minimum separability distance as defined in Spiess (1990). In the context of on-demand web mapping where the minimum separability distances are more severe (owing to the coarse display resolution) the depiction of buildings should be kept very simple. The concept here is also to define each new object as rectangle, whereby the area  $\overline{A}_n$  is computed out of the average of the pertaining buildings  $\sum A_k / nb$  in LOD<sub>25</sub>. As a consequence of the necessity to keep the minimal dimensions of buildings a scaling factor  $f_{area}$  is used:

$$\overline{A}_n = \frac{\sum_k A_k}{nb} \cdot f_{area} = \frac{\sum_k A_k}{nb} \cdot \underbrace{\frac{m_r^2}{m_{25}^2} \cdot \left( 1.0 - \Psi \cdot \frac{m_r - m_{25}}{m_{200} - m_{25}} \right)}_{f_{area}}$$

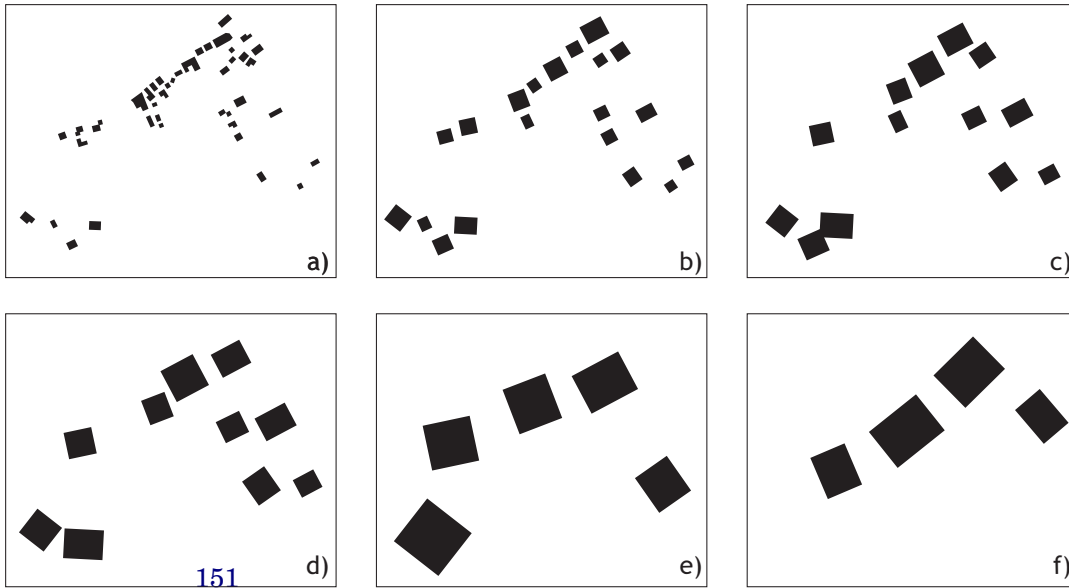
The first term  $m_r^2/m_{25}^2$  of the scaling factor compensates the reduction in size while the scale changes. Without the second part the buildings size would be always like in LOD<sub>25</sub> independent from scale. The second part is used to consider the size of buildings in LOD<sub>200</sub>. Their influence increases when the requested scale is closer to LOD<sub>200</sub>. Parameter  $\Psi$  has to be calculated from the ratio between original building size  $\overline{A}_{LOD_{200}}$  and calculated building size  $\overline{A}_n$  for LOD<sub>200</sub>:

$$\Psi = 1.0 - \frac{m_{25}^2}{m_{200}^2} \cdot \frac{\overline{A}_{LOD_{200}}}{\overline{A}_n}$$

Out of this corrected value  $\overline{A}_n$  the *width* and *length* of the new object can be obtained. The ratio *width/length* of the new building must be the same as for the largest represented object in LOD<sub>25</sub>.

**Orientation $\alpha$**  For the orientation  $\alpha$  of the new building only the value of the largest object of the group is considered. The reason for this approach is that the orientation of the largest building can be assumed to be most representative for the represented objects and influence or even dominate the characteristics of its environment. Hence, the orientation  $\alpha$  of the new and the largest object are equal.

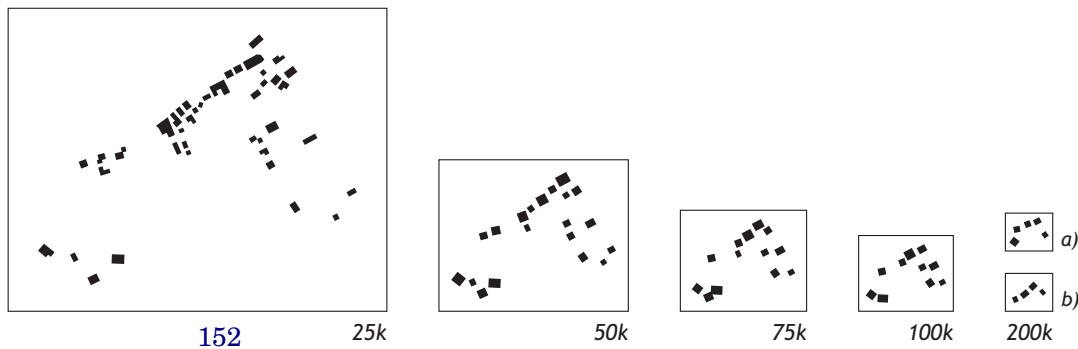
The Fig. 8.11a) shows the original data set of LOD<sub>25</sub> with  $nb = 41$  objects. For scale 1:50'000 (Fig. 8.11b)) the amount of

**Figure 8.11**

The positions and dimensions of the placeholders for different scales: a) 1:25'000 (representing LOD<sub>25</sub>), b) 1:50'000, c) 1:75'000, d) 1:100'000, e) 1:200'000 and f) 1:200'000 (representing LOD<sub>200</sub>) (all at scale 1:25'000). Data: VECTOR25/200 © swisstopo (BA034957).

objects decreases to  $nb = 20$ . The dimension  $[width, length]$  of each object depends on the dimensions of the represented objects, while the orientation is fixed by the same orientation as the largest building. In Fig. 8.11c)  $nb = 13$  objects are represented for scale 1:75'000. As can be seen overlap problems arise in the lower part of the figure. As one possible solution the elimination of one object can be taken into consideration. The Fig. 8.11d) shows  $nb = 10$  objects for scale 1:100'000. The last two figures 8.11e) and f) display the buildings objects for scale 1:200'000: e) the computed objects and f) the data set of LOD<sub>200</sub>. Comparing the size of the buildings at scale 1:200'000 a similarity can be assessed. Note that for all examples shown the constraints of minimal distance between individual objects has not been taken in account. All illustrations are depicted at 1:25'000.

The Fig. 8.12 shows the same situation, but now at the corresponding target scale. The Fig. 8.13 gives a complete example with the maps from LOD<sub>25</sub> and LOD<sub>200</sub> for comparison as well as the calculated position of placeholders and the resulting map for scale 1:75'000.

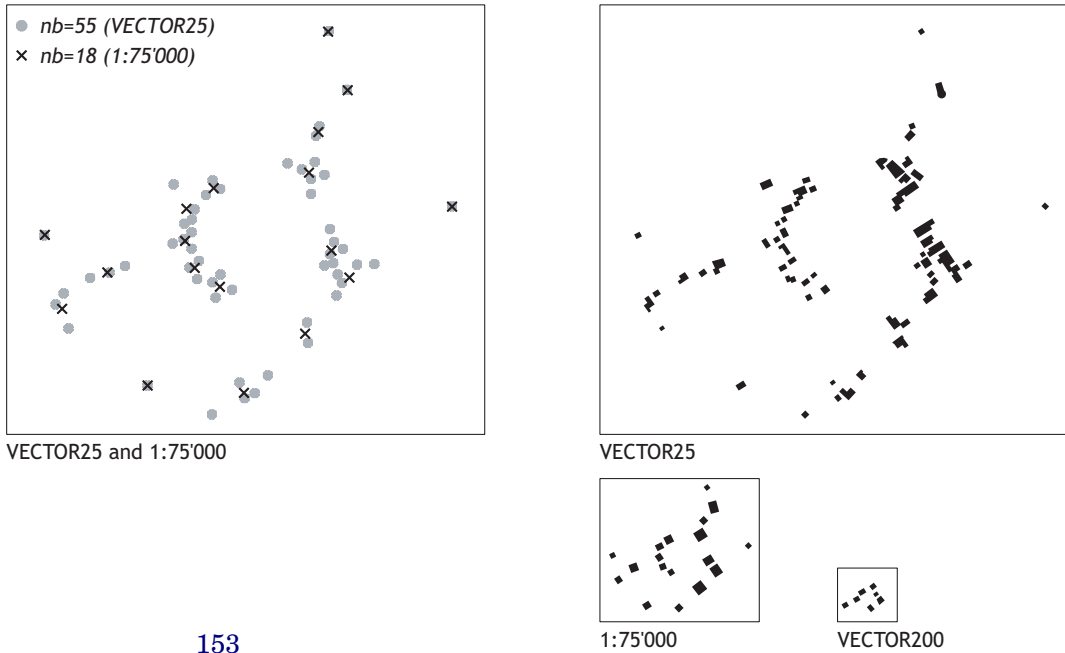


**Figure 8.12**  
Typification of  
buildings computed  
for different scales  
with correct  
proportions. For scale  
1:200'000 the  
computed (a)) and  
the original (b))  
representations are  
shown. Data:  
VECTOR25/200 ©  
swisstopo  
(BA034957).

## 8.4 Discussion of results and possible improvements

The typification of buildings is realized as two step process with positioning and representation. For the positioning step the mesh simplification approach was successfully applied. To determine the number of objects for the target scale independent from representation phase Töpfer's radical law was used. If the number of objects is known for several scales as given in Multi-Representation Databases an interpolation can be carried out to calculate intermediate scales. The proposed interpolation is based on an extension of the radical law with the number of objects at two source scales (e.g. 1:25k and 1:200k) and the parameter  $1/p$  as power, which influences the increase of the interpolation curve. With this modification the number of object can be adapted to manually generalized maps.

Advantages of using a Delaunay triangulation data structure during positioning phase is the fast calculation of distances between neighbourhood buildings. The weighting of the Delaunay edges is important for the control of building density. With the possibility of defining the correction factor  $s_u$  in the range of  $[0.0, 1.0]$  the user can influence the kind of map he/she wants to have. The differences between the representations concerning the positions of the building objects for ( $s_u \rightarrow 0.0$ ) or ( $s_u \rightarrow 1.0$ ) are significant. The advantage of using this method of mesh simplification is that any scale can be generated out of the two border data sets  $LOD_{25}$  and  $LOD_{200}$ . If the number of buildings is applied for the weighting the thinning process in densely populated areas can be decreased as shown in the examples. Furthermore object characteristics such as size, shape, orientation and semantics are considered as weights. The example showed that a weighting with

**Figure 8.13**

*Left: The positions of the placeholders for scale 1:75'000 (not to scale). Right: The dimensions of the placeholder computed out for scale 1:75'000.*

object size prevents a removal of large objects. A limitation during positioning phase is neglecting the alignments. Such particular patterns could be preserved by introducing additional constraints to the method. Constraints could be generated by an off-line pre-processing step that detects buildings alignments (for instance, using the method proposed by [Christophe and Ruas \(2002\)](#) and observed in the algorithm by setting increased weights for buildings in alignment structures.

Further improvements are needed for the representation step. The computed area of a placeholder depends on the areas of all represented buildings and is derived from the average of the area of all included objects. Even if a placeholder represents buildings with extremely different sizes the average size is displayed. This can influence the local nature of these few objects. A similar problem arises for example if one large-sized building and several small ones are used to compute the area of a placeholder. In this case the size of the new object will be too small compared to the large building. This situation leads to an unfortunate depiction of the placeholder. To better determine the size of a placeholder a statistical evaluation of the represented building objects can be accomplished.

## 8.5 Conclusions

The presented approach of mesh simplification is simple to implement and delivers reproducible results, which means running the algorithm twice on a set of building objects ends with the same unique solution. The whole typification process is divided into two steps, so-called position and representation. While the first step determine the number and the position of the new buildings with respect to the requested scale, the second step creates representative shapes at the determined positions. It was shown that object density could be controlled to emphasize areas with a higher building concentration and that object characteristics, such as size, shape, orientation and semantics can constrain the typification process. The approach allows the derivation of building objects for any intermediate scales with help of two fixed level of detail. The usage of a Delaunay triangulation data structures supports fast access during positioning step and therefore holds potential for real-time applications. The paper ends by considering the limitations of the suggested approach and shows tasks for further research and development.



## 9 Geo-enabling Spatially Relevant Data for Mobile Information Use and Visualisation

Alistair Edwardes, Dirk Burghardt, Eduardo Dias,  
Ross Purves and Robert Weibel

published in

Lecture Notes in Computer Science. 5<sup>th</sup> International  
Workshop on Web and Wireless Geographical Information  
Systems, Springer-Verlag, 2005, pp. 78-92

**Abstract.** This paper addresses the methodological issues of developing and visualising added value geo-enabled content for mobile information systems. The research was carried out under the framework of WebPark, an EC-IST R&D project that developed a location aware application for nature/protected areas. The evaluation of existing information sources - tourism information, research data, and multimedia content - revealed that the tourism information and multimedia content analysed did not have an active geographic component and the geographic research data (e.g. animal counts/observations) had a clear mismatch as regards the visitors' information needs. Therefore, different types of data processing were developed and performed in order to render existing information sources useful for location-based services. Data preparation also included building hierarchical data structures (quadtrees and hierarchical stream ordering). The paper shows how these data structures are exploited to facilitate real-time generalisation in order to efficiently present thematic point data on portable devices. Finally, three key lessons for geo-enabling location-based services are presented.

**Keywords.** mobile information services, LBS, environmental information, geo-enabling, location modelling, visualisation, hierarchical data structures, tessellations, real-time cartographic generalisation

## 9.1 Introduction

Protected areas are created, in the main, with the primary aim of conserving natural and cultural heritage and secondly in order to support the leisure and tourism industry. In this regard, delivering environmental education is a major task for most protected areas (Dias et al., 2004). Mobile technology is becoming increasingly available and its usage more and more widespread (Barnes, 2002) and mobile internet devices with geo-location capabilities may create an opportunity for delivering such education and information to visitors of protected areas. Location-based services (LBS) deliver information specifically customised to the user's location and, on occasion, context of use (Abderhalden and Krug, 2003) and as such can play a role in helping visitors to protected areas (e.g. national or regional parks) achieve full awareness of the richness of the surrounding natural and cultural re-sources and contribute to environmentally friendly visits. However, although the technology is mature enough to deploy mobile and context-aware applications, much of the data available is not ready to be used in such applications.

In this paper we address a set of issues encountered in making diverse and heterogeneous data resources accessible as value-added, geo-enabled content for mobile information services. These entail a range of challenges both in terms of how geo-graphic information is modelled and represented and, how it can be best be delivered to enhance exploration and understanding.

## 9.2 Context of research

This paper builds on experiences made in WebPark, a European IST project that developed a series of location-based services for users of protected areas (for information and sample pictures of the system, see <http://www.webparkservices.info>). The services are based on wireless technology and available for mobile smart phones and PDAs. The prototype system was developed at two partner sites, the Swiss National Park (SNP) and the Wadden Sea Islands (NL), and tested with real visitors (Abderhalden and Krug, 2003). The system includes the following features:

- ❑ automatic location of a user on a map (via GPS);
- ❑ points of interest (POI) search, both spatially and using free text;
- ❑ access to fauna and flora information, including multimedia data;

- ❑ context-aware retrieval of information, respecting visitors' spatial behaviour and personal preferences;
- ❑ dynamic mapping of retrieved information including real-time map generalisation;
- ❑ the ability for visitors to add their own geographic data as 'geographical bookmarks'; and
- ❑ location-sensitive warnings (e.g. areas with increased protection level).

The system uses wireless communication between the server and mobile devices (PDAs) with client side caching ensuring a continuous visitor experience when coverage is lost.

### 9.3 Evaluation of environmental and geographic information

Many authorities for protected areas are responsible for the collection and dissemination of information about their natural and cultural assets. Mobile information services can support such responsibilities by providing channels through which to access data holdings in geographically meaningful ways. Whilst the technology to allow this is relatively advanced, a problem remains in that most data resources held by such organisations are generally not in a state in which they can be made readily available. The reasons for this are well illustrated through the framework of [Raper et al. \(2002\)](#) for evaluating geographic information (GI). This model distinguishes between components of GI that are representational, at the levels of ontology, modelling and system, and communicative, at the levels of relevance, exploration, commodification and management. Table 9.1 uses this framework to describe the issues related to existing data resources encountered in the WebPark project.

Addressing these issues called for a series of information processing and management strategies. Here, we focus specifically on those related to data transformations.

### 9.4 Overview of data transformation and management

The WebPark system and related processes can be simplistically viewed as a publishing tool that allows intense information sharing of local knowledge with visitors to the park. In order to provide

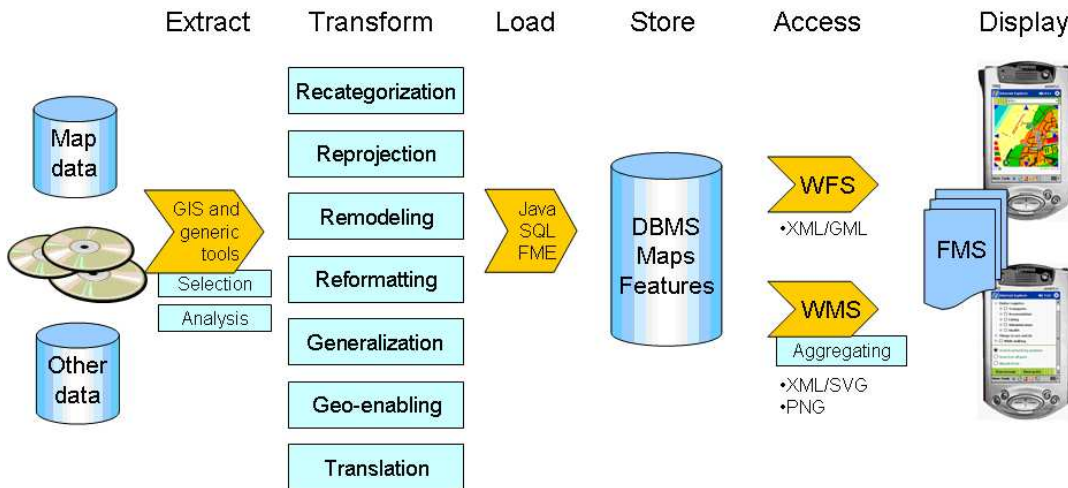
**Table 9.1**  
*Data considerations within the evaluation framework of Raper et al. (2002).*

Component	Level	Issues for available data holdings
Representational	Ontological	Entities were represented within data holdings using differing abstractions, classifications and taxonomies for similar or related information.
	Modelling	Diverse spatial and aspatial (e.g. multimedia) data models were employed amongst different data resources, with few or no geo-referencing relationships between them.
	System	A wide variety of data formats and media were used to store data (e.g. CD-ROMs, databases, documents).
Communicative	Relevance	Data had been collected for assorted dedicated purposes, for example research or inventory, and were generally mismatched to the information needs of ordinary LBS users Fehler! Verweisquelle konnte nicht gefunden werden..
	Exploration	Data were not structured or indexed in ways that made them easily adaptable to location- or other context based exploration. Data models were ill-suited for supporting readability, comprehension and visualisation, for example through multiple representations, generalizations or hierarchical abstractions
	Commodification	Data were not organised in chunks easily distinguishable as information products. Copyright issues related to using data within information services had usually not been considered.
	Management	Data holdings were largely decentralised and often purchased or produced for one-off purposes making management and update difficult.

true added value for the visitors, the information available plays a crucial role. Hence, GI and multimedia content needed to be adapted or created in order to meet the required accuracy and expectations of the visitor.

The GI content needed for the WebPark service can be divided into 'background' and 'foreground' information. Typical background GI consists of topographic base map data e.g. roads, paths, coastlines, water features and boundaries; false colour imagery classified by land cover; terrain information and public service and safety information. By contrast, foreground GI contains processed and interpreted GI and multimedia such as animal distributions, POIs, flowers in blossom and up-to-date photographs and other multimedia information.

In order to prepare GI content for WebPark, an extension to the ETL (Extract, Transform, Load) process from Vassiliadis et al. (2002) was defined. Fig. 9.1 illustrates the extended process.



**Figure 9.1**  
Data processing flow  
in WebPark.

The workflow can be considered as a linear process starting with the determination of available datasets (information audit) and ending with the display of information to visitors through extraction of data from its original storage medium, transformation to cope with the specified data model, and loading into a relational database. The extraction step also involves the analysis of the data in order to select a spatially and semantically relevant subset. The selected data is then transformed through a set of operations necessary to harmonise and standardise the data according to the WebPark data model. Depending on the individual dataset's these will include:

- ❑ Recategorisation - harmonising the conceptualisation of entities into a single do-main ontology.
- ❑ Reprojection - reprojecting spatial data into a common spatial reference system.
- ❑ Remodelling - mapping the different sources' data models into a WebPark one that allows access in a context- and location-sensitive way.
- ❑ Reformatting - converting the heterogeneous data formats (mime-types) for multi-media native content into the formats understood by the WebPark components.
- ❑ (Model) generalisation - applying geometric processes such as line filtering, aggregating features that are too small or defined with semantics that are too detailed.

- ❑ Geo-enabling - creating associations between spatial and non-spatial data
- ❑ Translation - Information is in English, French, German, Italian and Dutch.

Most of these steps can be easily mapped to the framework of [Raper et al. \(2002\)](#) shown in Table 9.1. For example, model generalisation improves readability and comprehension of mapping and thus can be considered as improving the communicative experience of exploration.

The processing steps that are methodologically most demanding and interesting are recategorisation, generalisation, remodelling and geo-enabling. Recategorisation is beyond the scope of this paper, but is described in [Edwardes et al. \(2003b\)](#). (Model) generalisation entails filtering and aggregating geometric as well as semantic information detail in a controlled fashion ([Weibel, 1997](#)). Different levels of detail are generated for the background data that can be associated with different target map scales. As the required procedures are relatively standard in GIS, they will not be discussed further. For a similar approach, see [Cecconi and Galanda \(2002\)](#). Remodelling and geo-enabling will be described in detail in sections 9.5 and 9.6, respectively. For more details on the other processing steps, see [Dias and Edwardes \(2005\)](#).

## 9.5 Remodelling

LBS primarily communicate information through two modes ([Edwardes et al., 2003b](#)); by answering questions about phenomena related to a user's context ([Abderhalden and Krug, 2003](#)) and by answering questions about the contexts of phenomena. Questions about a user's context (e.g. location and time of year) might be termed what questions, for example "What animals can be found here at the moment?" Those about the contexts of other phenomena might be termed where or when questions, for example "Where can eagles be found?" Questions relating to the phenomena, such as, "Why do we see more eagles near mountain ridges?", are answered secondarily, by attributing the phenomena (e.g. with multimedia) and by providing inter-connections between them, for instance with hyperlinks, map layering or detailed ontological models ([Agarwal, 2005](#)).

Depending on the nature of the phenomenon, answering the two types of primary question can require different spatial data models in order to relate it to the user's context and for presentation of spatial distribution. Using the terminology of [MacEachren](#)

(1994a, pp. 59-60), for discrete and particularly abrupt phenomena such as points-of-interest or real-time incidents, a single point-based model can be used, for example using distance to relate the user's location to the phenomena. If the instances of the phenomena are numerous, for example where they are discrete and smooth, presenting them may also require real-time generalisation (described in section 9.7 using the example of deer observation data).

More continuous phenomena, such as density of animal observations, exhibiting relatively complex patterns of spatial variation may need separate models with areas being used to relate the phenomena to the locations in the database (Jose et al., 2003) and a different cartographic model for presentation. In these cases, the cartographic content tends to be relatively static, e.g. seasonal animal distributions, and can be treated in a similar way to background data, using precomputed model generalisation to control data volume and the level of geometric and semantic detail.

Remodelling thus seeks on the one hand, to unify the diverse data models of the various data sources and, on the other, to restructure the data into models that can be used to answer the two primary types of questions. Such operations can be considered to first require consideration of the modelling of the data in the context of representation, and secondly to enhance the possibilities for exploration in the context of communication, as described in Table 9.1.

This section mainly deals with remodelling to relate the spatial occurrence of different phenomena to the location of a user. A process termed 'location modelling' (Schlieder et al., 2001).

### 9.5.1 Remodelling wildlife information

The main purpose of remodelling in WebPark was to allow a user to pose wildlife questions about their current situation. Achieving this requires a model of the user's location to which the spatial occurrences of plants and animals can be related.

It is possible to model location in different ways; geometrically, symbolically and semantically (Hu and Lee, 2004). WebPark used a geometric model where locations could be the user's current position or locations that could be indexed with respect to their current position such as topographic features and the spaces over which activities are performed, e.g. visibility regions and navigable regions. Wildlife phenomena were then associated to these locations.

Data about spatial occurrences of wildlife were found to have been captured in a variety of ways and using different types of geometry. For example, they might relate the entity to the geography influencing its spatial variation or behaviour, e.g. home ranges or



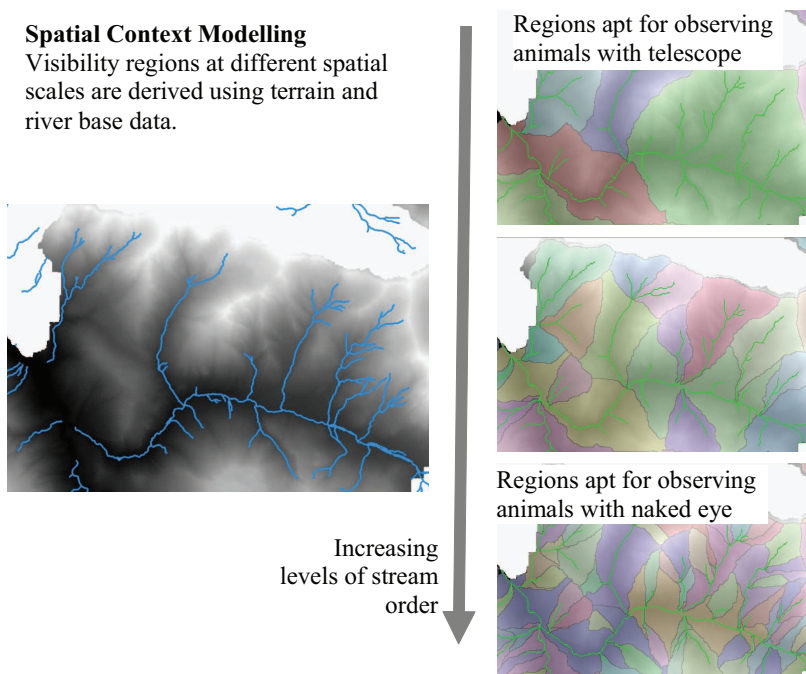
habitats, or they might relate them to the units used for recording direct experiences e.g. field observations, sections of path, census areas or areas delimited by expert knowledge. Mapping the phenomena onto the location model is thus essentially an intersection operation between the different geometry types. Table 9.1 illustrates this using a simple model of vector data intersections. These operations are performed either at query time in the case of point-point associations or the association is stored through an index related to the relevant object.

**Table 9.2**  
*Associating models of  
location and spatial  
occurrence  
geometrically.*

Phenomena		Point e.g. observation	Line e.g. river section	Area e.g. habitat zone
User				
Point	e.g. user's position	Distance, proximity	Buffer intersection	Containment and zone adjacency
Line	e.g. path	Buffer intersection or shortest path	Incidence or Hausdorff proximity	Intersection
Area	e.g. visibility region	Containment and region hierarchy	Intersection	Overlap or intersection

The approach implies that the location model can be defined completely independent of the phenomena, with the advantage that the time needed to search for information relevant to a user's location is considerably reduced. However, this assumption is in fact too restrictive. To some extent, the user's context is created both through their interaction with the phenomena they are interested in (Dourish, 2004) and to the size and behaviour of the entity. For example, large animals are usually very wary of people but their size means they can be observed from some distance using a telescope or binoculars; hence a location model of visibility can consider large regions. Small birds, even if tame, can only be observed at quite close proximity. This necessitates a model of visibility covering a much smaller extent. Figure 9.2 illustrates the generation of areas representing visibility regions in the mountainous terrain of the SNP for observing large animals such as ungulates. This involved computing the drainage morphology for stream channels at different stream orders (Strahler, 1952; Summerfield, 1991). The resulting hierarchy of regions gives an approximation of the area visible from a point at different scales. The approach has similarities with that of Schlieder et al. (2001) who also considering a hierarchical view of space for location modelling. The visibility regions were subsequently intersected with forest stands to take into account visual obstruction from vegetation.





**Figure 9.2**  
*Hierarchical stream basins and visibility regions.*

A similar operation was carried out to associate path segments with song bird occurrence. A habitat preference model for song birds (Filli et al., 2000) was overlain with buffered path segments. Each path segment was then attributed with song birds whose habitat preference intersected with their visibility buffer.

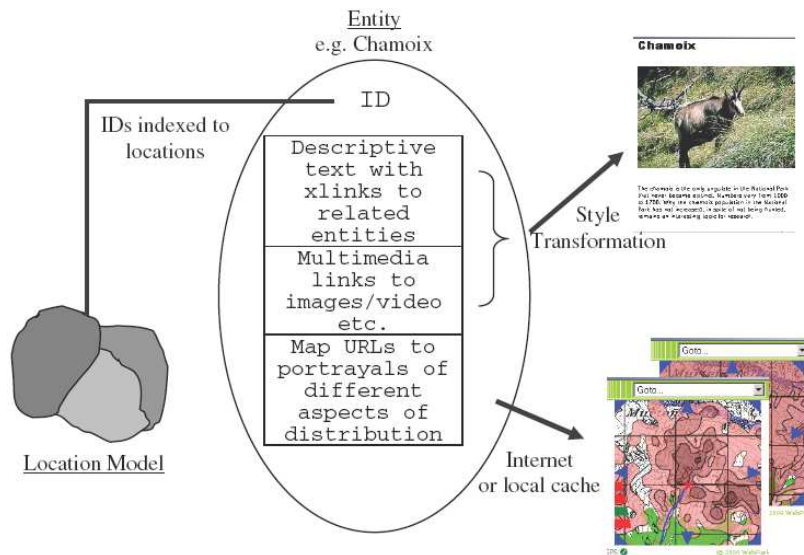
## 9.6 Geo-enabling

One of the main conclusions of the initial information audit was that there was a clear mismatch between the information currently provided by the park and the visitor information needs (Abderhalden and Krug, 2003; Dias et al., 2004). Most of the questions from visitors had a high spatial component (e.g. "What animals can I see around me?", "Can I have a picnic here?") but much of the most suitable information (e.g. the multimedia and documentation) was not geo-referenced. Information with a spatial component had generally been captured for research purposes and was not structured in a way that was relevant to the visitors' information needs. A fundamental issue was how to make these data fit for the purpose in an LBS. Geo-enabling is therefore an example of improving the relevance of information communicated to users within the context of Table 9.1.

### 9.6.1 Geo-enabling multimedia

Various sources of information, well fitted to visitors needs, were available as mul-timedia in digital form. For example, the SNP produces an interactive CD-ROM containing over 800 high-quality photographs as well as detailed texts, videos and sounds. Part of the richness of these resources is the numerous hyperlinks amongst the different pieces of information providing the ability for users to explore questions related to particular phenomena. A particular challenge was to design a model that would both ensure information was spatially explorable and that these inter-connections amongst pieces of information were maintained. Researchers on the GUIDE project (Cheverst et al., 2000) identified a similar problem. Their solution was to support separate information models for locations, containing places of interest and connections to travel between them, and for media, containing hyperlinked cultural and historical multimedia. The two models were then linked through the identities of places they had in common. This solution worked well for them because their primary interest was with places, with the system guiding their users between these. Thus their location model and entity ontology could to large degree be treated as the same thing. Our problem differed in that the entities of interest were wildlife and not inherently spatial phenomena. Hence different representations were need for entities and their locations. Our own solution was therefore to directly enrich the model of wildlife entities, with hyperlinks referencing the identities of related entities. Therefore, the entities became the main data chunks, around which other services could be commodified (see Table 9.1). This model is illustrated in Fig. 9.3.

The first step in this work was to recategorise data using an ontological mapping, such that, for example information about "bearded vultures" and information about "bartgeyer" were considered to relate to the same feature. The second step was associating these features with appropriate locations as described previously - thus for example bearded vultures might be associated with a path in terms of an area over which they were visible, but also with a point location used as a release site for a reintroduction programme. Finally, the entity description was enriched with links to appropriate multimedia and with links to associated features. Thus on querying for "What is near me?", a user might be informed that golden eagles (entity) and marmots (entity) were often seen in this area. On clicking on information about marmots, point locations of individual marmot sightings (location) could be displayed. Associated with these marmot sightings are multimedia describing marmots, and since golden eagles predate on marmots a link to the golden eagle entity. Using this model it was possible to have entities that



**Figure 9.3**  
Information model for  
WebPark.

were; non-geographic but still available to through system (e.g. general rules of the park, history), semi-geographic being spatially indexed to locations according to expert knowledge but having no mappable distributions, and geographic being both locatable and mapped.

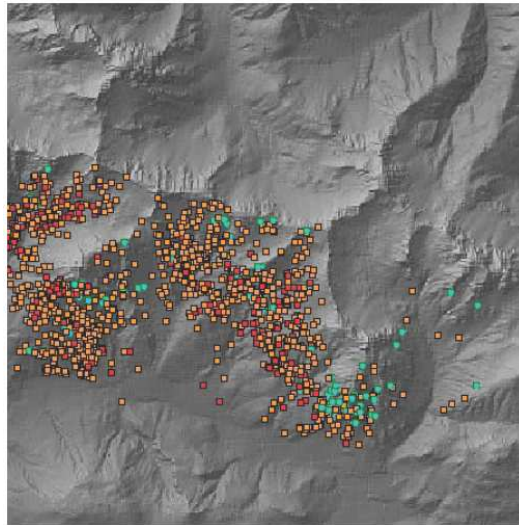
## 9.7 Design considerations for visualisation

In LBS information reduction is an important issue, for three main reasons. First, network bandwidth and client computational power impose limitations on the amount of information that can usefully be transmitted and processed. Second, the small screen size and resolution of portable devices imposes restrictions on the information content and density that can be portrayed in order to maintain the readability of visualisations. Thirdly, user interactions can result in rapidly changing portrayals (and associated levels of detail), for example through ad hoc querying or changes in location.

The first limitation is addressed in WebPark through a variety of techniques to minimise transmission of extraneous data. For example, where mapping is independent of location and user interaction is limited, e.g. topographic background maps, model generalisation as previously described is used in conjunction with client-side caching.

Developing methods to counter the second and third limitations is more demanding. While the detail of the (static) background data layers can be tailored to the small screen size during model generalisation and stored in a multiresolution database containing the right levels of detail for the target map scales, the portrayal of foreground data cannot be pre-computed and requires client-side, real-time processing since the foreground thematic data which is to be visualised is usually retrieved as the result of a specific ad hoc query and therefore can not be known in advance and pre-compiled in a map. Methods addressing these issues have at their core attempts to enhance explorative communication by support readability and comprehension through, potentially, multiple representations (Table 9.1).

**Figure 9.4**  
*Individual ungulate observations.*



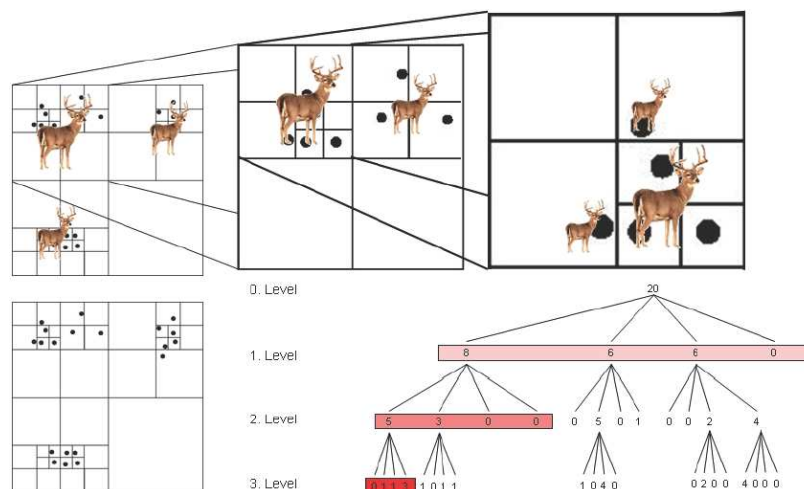
One example of data requiring such an approach are smoothly varying, discrete data such as those described in section 9.5. For instance, ungulate observations as shown in Fig. 9.4 are numerous and vary rapidly in density over space. Any representation of such points should meet three underlying constraints:

- ❑ the visualisations should be cartographically acceptable;
- ❑ any generalisation operations should be performed in 'real-time', that is, the user's wait for a response should be minimal; and
- ❑ the visualisations should maintain meaningful spatial and thematic relationships.

Since the data we are dealing with are simple point data, a minimal set of cartographic constraints is sufficient to maximize legibility on small devices: first the minimum symbol size should maintain legibility (a common value is 5 pixels, [SSC \(2005\)](#)); and second the symbols should not be allowed to overlap or touch.

### 9.7.1 Using spatial tessellations for real time map generalisation

Many approaches exist to generalizing and transforming point data ([de Berg et al., 2004](#); [O'Sullivan and Unwin, 2002](#)). Our approach to this problem is based on the notion of hierarchical spatial tessellations. Quadrees ([Samet, 1990](#)) have been widely used in both computer graphics ([Berman et al., 1994](#)) and GIS ([Timpf, 1997](#)) to index and allow rapid traversal and retrieval of data. Here we use the quadtree to make decisions on the number of objects to display. Moving vertically through a quadtree can be considered analogous to zooming, whilst moving horizontally between nodes is comparable to panning. Each node contains a count of the number of objects existing within the block it represents. The count at the nodes for the current zoom level is used to generate a visualisation using simple selection and exaggeration rules (Fig. 9.5, lower part). The quadtree tessellates space until every point is assigned to a separate block. When zooming a level is chosen that meets a minimum acceptable symbol size criterion.



**Figure 9.5**  
Lower part: Example point set and corresponding quadtree (note the count stored at each node). Upper part: Successive generalisations of deer observations.

The upper part of Fig. 9.5 illustrates the results of zooming/generalisation through the quadtree, where maximum symbol

size is constant, and symbol size is a function of the number of observations. Symbols are also displaced, within the quadtree block, toward the centre of gravity of the block's observations.

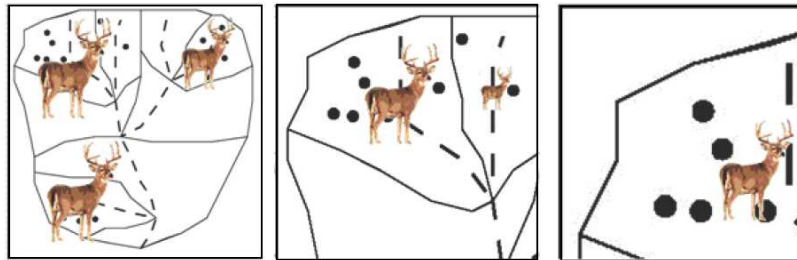
The quadtree provides an efficient implementation which allows the application of simple rules to generalize thematic point data for panning and zooming whilst meeting basic cartographic constraints. However, points are aggregated according to an arbitrary quadtree block, with no consideration given to the distribution of points either with respect to other points (i.e. point density) or relationships with other underlying variables (i.e. the position of a point within a landscape) (O'Sullivan and Unwin, 2002).

### 9.7.2 Using catchment order in tessellation

In Section 9.6 the use of hierarchically ordered catchment areas, based on stream orders to spatially index locational data was discussed. The main reason for using such an index was a hypothesised link between visibility and catchments.

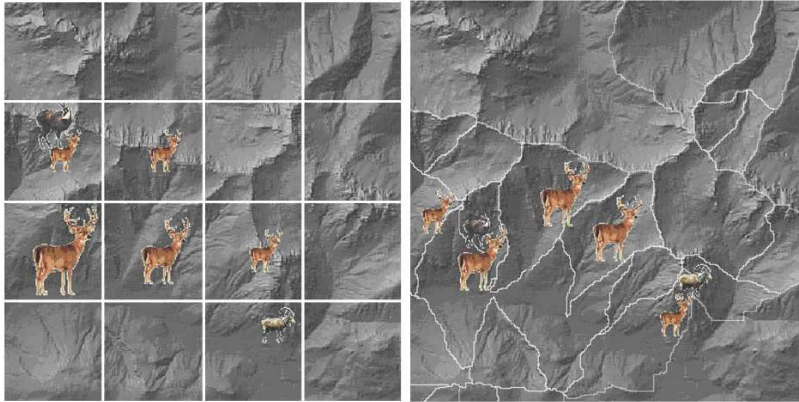
Here we postulate a relationship between deer and the landscape. Deer in mountainous terrain may be unable or have difficulty to cross steep ridges, and thus individuals in one valley may be more likely to have some relationship with each other than spatially nearer deer in separate valleys. We therefore once again tessellate the point dataset in Fig. 9.5 using Strahler ordering and watersheds (Summerfield, 1991), but do not on this occasion time apply the mask of forested areas.

**Figure 9.6**  
*Idealised Stahler stream ordering and tessellation using catchments at three levels showing progressive levels of zoom.*



An idealised example of such a tessellation is shown in Fig. 9.6. Fig. 9.7 shows a comparison between generalisations for the third level in a hierarchy based on the point data shown in Fig. 9.4. Here, the difference between the two images is striking and demonstrates the importance of using a geographically relevant unit of aggregation to tessellate space (Openshaw and Alvandies, 1999). It is important to point out, that in this case, the link between catchments and ungulates is hypothetical.





**Figure 9.7**  
Comparison of hierarchical generalisation techniques using quadtree (left) and hierarchical stream ordering and associated catchments (right).

## 9.8 Conclusions

This paper reported on lessons learned in making environmental data available through LBS, using the WebPark project as an exemplar. The paper focused on representation and communication of geographic information and some of the steps required in order to enhance these components of GI. In particular the paper described the steps necessary to model locational information and an approach to real time generalisation whilst maintaining geographic relationships.

Although the paper describes a single project, some important general lessons can be drawn, which have in general not been considered by the bulk of previous research focused on predominantly point based representations of the world (in an LBS context).

The first key lesson, described in section , is that environmental data are not commonly found in a form that is fit for the purpose of an LBS - a lesson with resonance to [Cheverst et al. \(2000\)](#). Typically, the data have been collected with an alternative purpose in mind and the process, and most importantly the effort required, to transform them for use in an LBS is non trivial.

The second lesson, encountered in sections 9.5 and 9.6, is that not all locational data and relationships are best represented as points. Dependent on the context and location of use multiple representations of the users' location and related phenomena may be appropriate, providing a rich and diverse set of potential query types and possibilities for interaction and understanding.

The final lesson, highlighted in section 9.7.2, is simple. Don't forget the geography. Environmental data rarely, if ever, exist in isolation and thus transformations must consider how to preserve geographically important relationships. This in turn requires us

to consider how best we can explore data in order to identify such relationships and requires support from domain experts.

## Acknowledgements

The authors gratefully acknowledge the support of the WebPark team (<http://www.webparkservices.info> - Geodan Mobile Solutions, EADS Systemes & Information, City University London, Laboratorio Nacional de Engenharia Civil, and the Swiss National Park), the European Commission for supporting the WebPark research (Project Number IST-2000-31041) and the Swiss State Secretariat of Education and Research (SER) in relation to this project (BBW Nr. 01.0187-1). Eduardo Dias would like to thank the support from the Portuguese National Science Foundation (FCT/MCT) under the PhD grant SFRH/BD/12758/2003. The WebPark system continues to be developed and marketed through a spin-off company Camineo (<http://www.camineo.com>).



## 10 Automated Generalisation of Dynamically Generated Maps for Mobile Devices

Dirk Burghardt, Alistair Edwardes, Ross Purves  
and Robert Weibel

published in  
Kartographische Nachrichten (originally in German),  
Vol. 55, No. 3, 2005, pp. 119-125

**Abstract.** Automated map creation has attracted significant attention because of the need for dynamic applications in web mapping and mobile information systems. Information to be delivered can be filtered with respect to user profiles, personal preferences or a user's current position. In these cases pre-calculation and storage of cartographic presentation is not a tractable solution. This article presents approaches for real time generalisation of dynamically generated maps in response to individual information requests, with a focus on cartographic presentation of symbolised point objects. Applied methods are clustering and the use of quadrees for typification, as well as hierarchical data structures for adaptive zooming and panning. Tessellations based on information derived from digital terrain models also allow a consideration of background information. The development of the presented approaches was carried out in the context of the EU project Web-Park whose aim was the implementation of a platform for delivering location-based services in recreational and protected areas. A result was the creation of a mobile information system which supports user orientation in unknown areas and allows searching of extensive thematic information resources.

**Keywords.** LBS, mobile information services, environmental information, hierarchical data structures, tessellations, real-time cartographic generalisation

## 10.1 Introduction

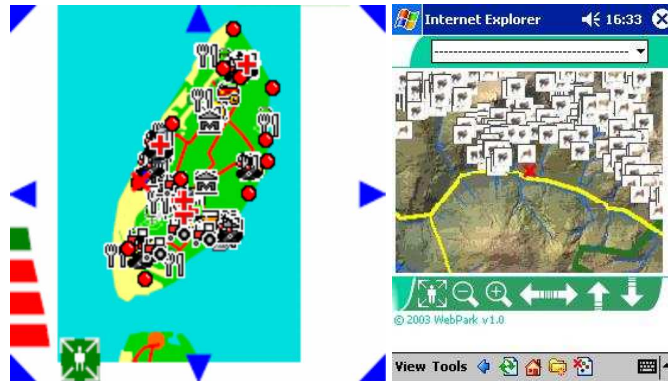
Research in the area of automated generalisation has concentrated in the past on the development of generalisation tools for topographic maps. Several main tasks were identified like conflict detection and conflict solution, as well as the utilisation of "cartographic constraints" for the description of cartographic acceptable visualisations without mapping conflicts (e.g. Ruas (1998); Weibel and Jones (1998)). The runtime of the algorithms did not receive much attention, because the organisation of map production could be adapted.

In contrast the presentation of topographic data over the internet is time critical. Early developments in the web mapping domain are mainly technically oriented, driven by, amongst others, the Open Geospatial Consortium, Inc. (OGC). Web Map Services support some user interaction as zooming and panning, mostly with precalculated topographic maps. Therefore a Multi-representation database is generally utilised. The first algorithms for automated generalisation in real-time were developed for mobile applications by Harrie et al. (2002), as well as for adaptive zooming from Cecconi and Galanda (2002). The disadvantages resulting from a small screen size can be compensated by user interaction.

New forms of mapping which allow the user to query not only individual layers of topographic data, but to ask questions of associated thematic data, provide new challenges in delivering cartographically appropriate representations. Here the thematic data which is to be visualised may have been retrieved from a database as a result of a specific user query and may be only one of many possible solutions which are dependent both on the user's personal preferences and location. Thus, precomputing and storing potential visualisations is not a tractable solution and solutions must be computed on-the-fly.

Figure 10.1 shows two negative map examples of ungeneralised maps with poor legibility. Symbols overlap each other and can not be separated from each other. Thematic maps traditionally display no more than a few variables and their variation over space, perhaps with a topographic backdrop (Slocum, 1999). Typical examples include maps portraying animal locations or election results by ward. Such maps aim to allow users to identify relationships within and between datasets as well as displaying the locations and values of attributes at absolute locations.

In this paper an approach for on-the-fly generalisation is presented, developed as part of the WebPark-project. The general aim of the project was the creation of a mobile information system for protected and recreational areas (Burghardt et al., 2003). Queries on thematic information were combined with location-based ser-



**Figure 10.1**  
*Ungeneralised thematic maps with places of interest (left) and animal observation data (right).*

vices. Examples are the presentation of interesting places, hiking trails, as well as other important objects (points of interest) with consideration of the current position of the user. Additionally, a service was designed for the selection of observation points for animals and plants.

Next a set of constraints is identified for the cartographic presentation of symbolised point objects on mobile devices. Following this, three different approaches will be proposed for the on-the-fly generalisation. Clustering approaches will be applied for the aggregation of point objects considering the relative position. The usage of a quadtree data structure supports adaptive zooming and panning. Furthermore an extended approach is shown, considering relationships between thematic foreground and topographic background information.

## 10.2 Constraints for generalisation of symbolised point objects

### 10.2.1 Cartographic constraints

The data dealt with are simple point data and a minimal set of cartographic constraints are considered, which are intended to maximise legibility on small devices:

- ❑ minimum symbol size should maintain legibility - a common value is 5 pixels (SGK, 2002); and
- ❑ symbols should not be allowed to overlap or touch.

Additionally the cartographic presentations for mobile devices should be generated dynamically, that is, the automated generali-

sation has to be realised on-the-fly. Therefore, a distinction is made between foreground and background information. Background information is mainly static and does not need to be updated very often. Examples are topographic information, borders, relief presentations and so on, which provides a spatial referencing (creating the base map). In contrast foreground objects represent thematic information, which will be queried by the user in arbitrary combinations. Additionally the information can be adapted to the user requirements, the current task, as well as the location and time (Reichenbacher, 2004). The cartographic presentation has to be dynamic and can not be generated completely in advance.

### 10.2.2 Maintaining important spatial relationships

The derivation of thematic presentations should also consider the relevant spatial relationships besides the above mentioned minimal requirements. O'Sullivan and Unwin (2002) distinguish between relations of first and second order. Relations of the first order mainly influence absolute positioning of foreground objects, because they model the association of fore- and background objects. Relations of second order describe interactions only between foreground objects and have therefore an effect on relative positioning.

For instance, in considering the positions of several deer, does may be clustered around a stag displaying such a second order relationship. However, if the deer are separated by some physical feature in the landscape then our cluster may have no biological meaning and thus not be relevant. Equally, deer may be found in areas where a preferred food source occurs e.g. salt stone and thus variation in density of deer over space is controlled by some other parameter displaying a first-order effect. De Berg et al. (2004) describe a wide range of algorithms which can be used to identify such relationships and simplify point maps. If we wish to preserve meaningful relationships in our thematic data we must find ways to consider both first and second-order effects.

## 10.3 Approaches for conflict free presentation of symbolised point objects

### 10.3.1 Removal of symbol overlapping by clustering approaches

The automatically generated maps of the mobile information system WebPark are based on observation data of animals and plants of the past years. The information content of the data consist

amongst others in the spatial distribution, characterised through different densities and the inherent patterns. The aim of the generalisation is the maintenance of this information. An often used generalisation operator applied therefore is the typification.

Our approach for typification is called hierarchical agglomerative clustering and is one of the standard methods from the information analysis domain. Starting from one cluster for every object a pair wise unification is applied, based on a comparison of similarity or distance measures. The method is carried out iteratively until a given number of clusters has been reached, until one cluster remains or the distance between the closest two clusters becomes longer than a given threshold. Applications of clustering approaches in generalisation can be found in [Regnauld \(1996\)](#), as well as [Anders and Sester \(2000\)](#). There a hierarchical graph based clustering is applied for the identification of building groups. A further reorganisation of the building groups is carried out with the displacement tool PUSH, described in [Sester \(2000\)](#).

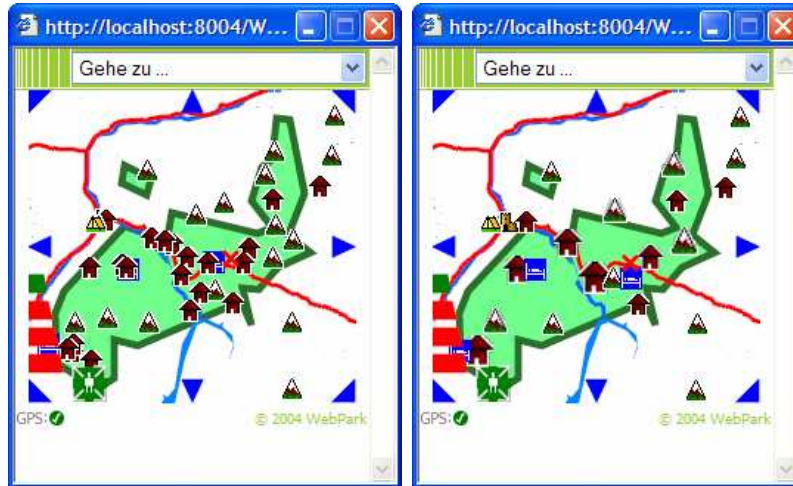
The procedure for the removal of symbol overlaps can be described as follows. The first step is the grouping of single object classes, for example huts. The criterion for unification is based on euclidian distance, i.e. in one iteration step the two closest symbols will be aggregated and presented with a new one. The position of the new symbol can be derived from the center of gravity, in the case of two symbols this is half the distance between them. In a second step the clustering is carried out with all object classes, for example huts, hotels and sights. Clusters of symbols from several object classes can not be aggregated, therefore symbol placement around the center of gravity with consideration of original positions is applied (Figure 10.2).

The aggregation of symbols is indicated by a larger symbol size. The size value of the new symbol grows non linearly with the number  $n$  of represented objects, whereby the symbol size can not grow in an unlimited way. The factor  $f$  for the enlargement of symbol width and height varies between 1 and a maximum value  $f_{max}$ . The calculation is done by

$$f = (f_{max} - 1) \cdot (1 - g^{(1-n)}) \quad (10.1)$$

where  $g$  is the control parameter for the increase of symbol size. Table 10.1 shows the increase factor  $f$  with dependence on the number  $n$  of represented objects. Larger values of the control parameter  $g$  also speed up the increase of symbol size. A maximum value for the enlargement was set to the double symbol size  $f_{max} = 2.0$ .

**Figure 10.2**  
Reduction of symbol overlaps with clustering approaches (left: presentation without generalisation, right: result of clustering).



**Table 10.1**  
Factor for enlargement of symbol width and height.

Number objects $n$ per symbol	$f$ for $g = 1.2$	$f$ for $g = 1.5$
1	1.0	1.0
2	1.17	1.34
3	1.31	1.56
4	1.42	1.70
5	1.52	1.80

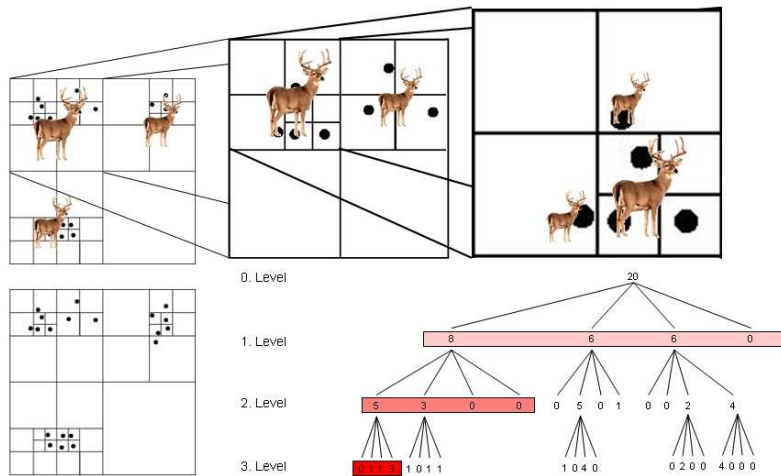
### 10.3.2 Hierarchical data structures for real time access on the example of quadtrees

Computing time is a fundamental criterion for the dynamic creation of maps on mobile devices. Small screen sizes require on the one hand a higher degree of abstraction and with that more effort on the model generalisation and information filtering. On the other hand the number of objects portrayed on the screen is comparatively small, which allows a cartographic generalisation to be performed in real time. A further simplification of the dynamic map creation can be reached through a distinction between foreground and background objects. While a presentation of background objects can be calculated previously, the foreground objects have to be placed dynamically at runtime. To this end the utilisation of hierarchical data structure is helpful.

The application of hierarchical data structures is well known from applications such as 3D landscape visualisation. There objects are presented with varying detail, so called levels of detail (LOD) dependent on the distance from the observer. Applications

of hierarchical data structure are already known in cartography and particularly in automated generalisation. Examples include the BLG-tree and the GAP-tree (van Oosterom, 1989, 1995) or the usage of reactive data structures for automated label placement (Petzold, 2003). Recent work by Neun et al. (2004) deals with the modelling of vertical relations inside multi-representation databases (MRDB). The usage of quadtrees for generalisation, especially for line simplification, was proposed the first time by Dutton (1996), though with a specific triangle based quadtree structure.

Usually a quadtree data structure is created by an axes parallel, recursive subdivision into four parts. A possible construction is realised with equal partition at the centre of the quadtree box. Figure 10.3 illustrates the application of quadtrees for adaptive zooming. The example shows the first four levels of a quadtree data structure. Already by the third level nearly all nodes are represented by leaves, which means a maximum of one object is contained. The number of nodes per level equals  $4^{level}$ , for example the third level contains a subdivision into 64 quadtree boxes. The geometric series  $(1 - 4^{level-1})/3$  enables the calculation of the entire number of nodes inside the data structure. For the case of the subdivision down to the third level 85 nodes are contained.



**Figure 10.3**

*Usage of quadtree data structure for adaptive zooming.*

The quadtree data structure can be applied for the real time creation of thematic map presentations. The necessary information is accessed from the relevant nodes dependent on the zoom level and selected map area. Moving vertically through the quadtree is directly analogous to zooming, whilst moving horizontally between nodes is comparable to panning. Each node contains a



count of the number of objects existing within the block it represents. The count at the nodes for the current zoom level is used to generate a visualisation using simple selection and exaggeration rules.

The size of screen and icons determine the number of symbols which can be presented side by side. For instance 8 x 8 boxes over a (PDA) screen size 240 x 320 pixels altogether 64 overlapping free boxes of size 30 x 40 pixels. The usage of quadtree provides a suitable data structure for real time generalisation based on regular space partitioning. A disadvantage of such a subdivision is noticeable in Figure 10.4 (middle), where the underlying grid structure becomes obvious. Alternatively symbols could be placed at the center of gravity derived from all observations of the quadtree cell. In that case, however, a visualisation without symbol overlaps can no longer be guaranteed. Also experiments with a reduction of the cell size might be possible (i.e. more quadtree levels), which would cause more computational effort.



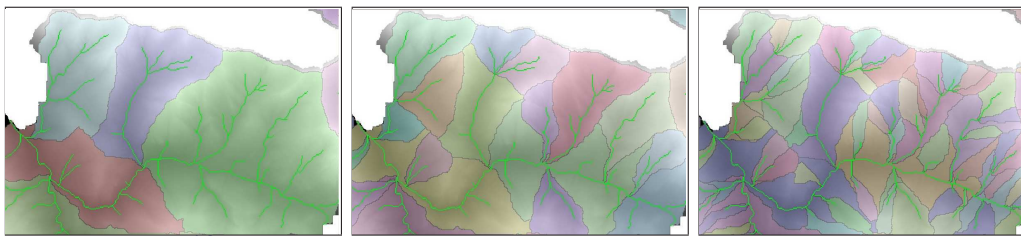
**Figure 10.4**  
Removal of symbol overlaps (left) with the quadtree approach (middle) and clustering methods (right).

The statements about hierarchical data structures on the example of quadtrees apply also to the typification based on clustering. In clustering a dendrogram is first constructed. The dendrogram, then, can visualise details as well as overview information depending on the position inside the dendrogram tree in a similar way like quadtrees.



### 10.3.3 Mosaic subdivision based on background information

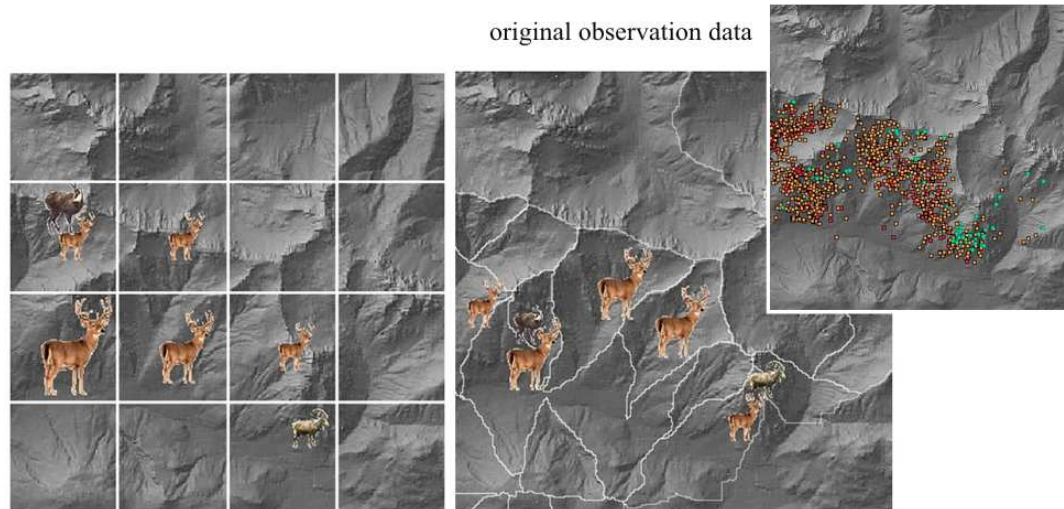
The presented clustering approach for the removal of symbol overlaps is suitable to preserve relative arrangements of point objects, i.e. second order relations (O'Sullivan and Unwin, 2002) at different scales. First order relations, representing associations between foreground and background, are neglected so far. Assumptions are necessary about the relations between foreground and background data. In our application context the visitors should have access to information which gives them a forecast for the probability of seeing different kinds of animals in a particular valley.



**Figure 10.5**  
*Mosaic subdivision on the basis of the digital terrain model and hierarchical stream basins (catchments).*

Therefore a hierarchical mosaic subdivision is applied on the basis of the digital terrain model of the study area. First of all the flow accumulation network is derived from the digital terrain model, which corresponds relatively well with the river network. Following that, the Strahler ordered network (Summerfield, 1991, pp. 208-209) was derived to find the appropriate catchments (Figure 10.5). Higher order tessellations use the catchments of that order as well as the lower order ones. Finally, by analogy with the usage of the quadtree data structure, the number of observations per animal species can be stored for the different catchments depending on the tessellation order (Burghardt et al., 2004b).

Figure 10.6 shows a comparison of a thematic presentation derived with the quadtree approach vs. the mosaic subdivision. For the mosaic subdivision approach catchment size has to be considered. Otherwise, if the catchments are too small, a conflict free symbol placement is impossible. A simple solution might be based on the comparison of symbol size and the maximum inscribed rectangle for the different catchments.



**Figure 10.6**  
Comparison of  
hierarchical  
generalisation  
techniques with the  
original observation  
data. Left: Using  
quadtree. Right:  
Using hierarchical  
stream ordering and  
associated stream  
basins.

## 10.4 WebPark project - mobile tourist guide for recreational and protected areas

The work presented here on real-time generalisation of dynamic generated maps has been part of a European research and development project, the main goal being to develop a mobile information system for recreational and protected areas. The focus of the project was application oriented, which meant two thirds of the time was spent on data preparation, user questioning, application development and performance improvements. The final result is a mobile walking guide, useable by tourist and visitors of several national parks.

### Developed services

- ☐ automated presentation of current user location and dynamic map creation on the basis of thematic information
- ☐ comprehensive search request with and without location filter (routes, mountain profiles, observation of animal and plant, as well as sights and points of interest)
- ☐ information ranking based on information analysis with intelligent agents ([Mountain et al., 2003](#))
- ☐ creation and request on geographic bookmarks

There are several location-based services accessible, for instance the selection of routes and point of interest, storing of geographic bookmarks as well as requests on extensive thematic information (Burghardt et al., 2003).



**Figure 10.7**  
*Group with mobile walking guide at the Swiss National Park.*

In the following some experiences are itemised, which are possibly characteristic for projects with a technological focus. The main part of the research and development work was needed for the adaption and creation of basis technology for the provision of location-based services. Tests at the beginning of the project showed that, especially in (alpine) recreational areas, network coverage is rare. To become independent local caching strategies were applied to guarantee access on information. A local caching http-server was developed for the client to handle requests at first on the PDA. In the case that the required information is not available in the memory cache, the request will be forwarded on to the server.

A difficulty came from the fact, that the client software had to be adapted several times, because the hardware of the PDA and smartphone changed rapidly and therefore nearly every year new devices were used. In addition input from continuous user testing needed to be integrated. Nevertheless, it was an advantage to start quite early with such prototype testing. That way the cooperation and communication between the developers was improved, as well as the user requirements being identified quickly and considered during the development.

Figure 10.8 shows the different development states of the graphical user interface. The picture left in figure 10.8 represents a early version (2003), the picture in the middle the current version of map view (2004) and the picture on the right shows the selec-

tion menu for the different services and search functions (2004). It can be seen for instance that during the projects the pan and zoom button were integrated in the map view, with the advantage of an increased mapping area. Furthermore, the user guidance was improved through the usage of icons to access frequently used applications. In Figure 10.8 right both possibilities are visible, the change of services via icons (in background) as well as the access through a drop-down-menu.



**Figure 10.8**

*Improvement of information access based on changes of the graphical user interface.*

The WebPark-project finished in October 2004. The EU commission rated the project as very successfully and listed them on reference projects of the Fifth Framework Programme. In addition, a spinoff company called Camineo was founded by the French partner involving the WebPark groups as shareholders.

The finalised platform can provide an excellent basis to investigate further research questions on the domain of mobile information systems. Of great interest, for example are analysing the process of information search and, from a cartographic point of view, investigating if there are particular adapted map presentations more suitable for dedicated user groups and applications. A further questions might concern the influence of context based technologies on peoples behaviour and their individual perceptions of the environment. In this sense it would be interesting to know if tourists using a mobile guides are really better informed than visitors accessing conventional information sources and also if the orientation task becomes easier.

# 11 Conclusion

## 11.1 Discussion of the papers presented

In Chapter 3 the current state of the art and a summary of potential new areas for research in generalisation was presented. In the following discussion, the work presented in this habilitation thesis is confronted in particular with upcoming research topics identified from the generalisation workshops summarised in Table 3.2. By summarising the research questions addressed by each paper, it will be shown that three of these four research themes are addressed in this thesis. Furthermore, key directions for further research will be suggested.

The first paper in this habilitation thesis addresses the research on *web generalisation services* (upcoming topic 4 presented in Table 3.2). The paper reviews the advantages of the web service concept and proposes a categorisation of generalisation services into support, operator and processing services. Furthermore, an architecture is described such that generalisation functionality can be offered independently from usage scenarios (e.g. on-the-fly mapping or tool chaining), hardware and software used and the programming languages applied for the implementation of generalisation algorithms. In the longer term, such approaches should allow the sharing of implemented generalisation algorithms and thus facilitate research in the dynamic and evolving area of geovisualisation and digital cartography.

*Chapter 4*  
(*first paper*)

Open research questions relate to for example semantic service description and discovery as well as service composition and orchestration. For example, ontology-based matching solutions (Das et al., 2004) are one possible route to exchange service descriptions on the schema level as they match geographic features across products. Further research should also investigate the degree to which service granularity is limited by the effort involved in exchanging and transforming data and furthermore, to categorise different support, operator and processing services (Neun et al., 2006; Burghardt and Neun, 2006).



Chapter 5  
(second paper)

The second paper discusses the utilisation of principal components analysis (PCA) in exploring relationships between cartographic constraints. As such this paper addresses research on the formalisation of cartographic constraints (topic II, Table 3.1) and the explicit modelling of spatial relationships (topic III, Table 3.1). The first part of the paper addresses the detection and modelling of building alignments. The second part applies the method to formally investigate correlations between cartographic constraints. Therefore a *constraint space model* is proposed providing a formalised description of cartographic constraint violation. Using such a formalisation it is possible to both analyse the current cartographic situation and to trigger and orchestrate generalisation operators. Finally the paper shows possible further applications of PCA to the generalisation process, for example in the detection of settlement types from building data sets (Steiniger et al., 2007).

Besides investigating correlations between constraints, it would be interesting to analyse the influence of the different generalisation operations as changes represented in the correlations between the constraints. Therefore, successive iterations in the generalisation process must be compared by PCA. Further research in the domain of constraint formalisation should explore suitable strategies for the selection of generalisation operators on the basis of constraints and investigate the weighting and prioritisation of constraints. A related problem is the handling of a large number of cartographic constraints - should they all be considered simultaneously or is it possible to separate independent generalisation subtasks described by a reduced number of constraints? Finally, the constraint model has the potential to control the map design process on a more general basis, including the selection of relevant information and important relations on the basis of map purpose, user requirements and context of use.

Chapter 6  
(third paper)

Paper three on explicitly modelled relations within a Multi-Representation Database extends research on constraints and the modelling of spatial relationships to a more general level. Underlying this work is the expectation that additional explicitly formalised information, that goes beyond the modelling of features with geometries and attributes, is needed to enable successful automated generalisation. In this paper three types of relations are investigated, namely horizontal, vertical and update relations. The relations are further defined as a characterisation of features and their changes across scales and time with generalisation relevant meta information. The research and development work presented in this paper thus belongs to topics III and V of the generalisation research categories presented in Table 3.1.

Open research questions emerge from the consideration of horizontal relations during the generalisation process. Firstly, they constrain the generalisation of the map features, mainly through requirements for preservation of the original state. Furthermore, from a hierarchical viewpoint the horizontal relations are themselves elements of the generalisation process, which can also be emphasised or typified. Currently, there are few approaches which consider horizontal relations explicitly. Examples include the preservation of topological relations or the generalisation of specific cartographic patterns such as building alignments and street bends. Further research is necessary on the usage of vertical relations to enable automated incremental updates. Depending on the number of changed features initial and incremental generalisation strategies may vary (Bobzien et al., 2005b). Additionally separation strategies are required to maintain changes locally. As well as the creation of relations between features, relations are also applicable at the schema level for the assignment of feature classes.

The subsequent papers four and five in this habilitation thesis describe new approaches to the implementation of generalisation algorithms (topic IV of Table 3.1). The two approaches presented here, smoothing by snakes and typification by mesh simplification, have some commonalities. Firstly, both snakes and mesh optimisation are based on optimisation approaches. The advantage of optimisation methods is that they support simultaneous application of different generalisation operations at the same time, avoiding the problem of operator sequencing. In addition both papers present solutions to the typification operation. While the snakes paper focuses on the typification of linear features such as road bends, in the mesh simplification paper a typification solution is proposed for point like features, such as buildings, labels or icons. This second approach in particular has potential to be applied for varying generalisation purposes such as the generalisation of topographic maps (e.g building typification) or the clustering of point like features such as icons on mobile maps.

*Chapter 7 and 8  
(fourth, fifth paper)*

The development of basic generalisation algorithms is ongoing across the generalisation community. A major area where progress can still be made is in the development of algorithms which can be applied in real-time or allow the derivation of hierarchical data structures supporting real-time access. This is exemplified by our algorithm for mesh simplification. Equally important for future developments in generalisation algorithms are the subsequent consideration of spatial and semantic relationships. Currently, although some algorithms support the preservation of topological relationships, very few deal with the explicit encoding of spatial patterns or semantic relations.



Chapter 9  
(sixth paper)

Paper six investigate generalisation tasks in the context of mobile information systems (upcoming topic 3 in Table 3.2). One major trend in this domain is reflected through the *shift from general to specialised maps*. For example, tourist information systems as developed in the WebPark project aim to provide information adapted to location and context. Furthermore, the function of general topographic maps is being reduced to the simple provision of background information overlain with user specified, location dependent or temporally relevant thematic information. In such applications, aerial photographs or visualisations of terrain compete with the topographic map to offer background information (e.g. Google Earth).

Related to this trend of delivering specialised information is a strong *user centred focus* on map content presentation. Methods to support such user centred focus, span the possibility of new modes of user interaction to completely automated applications providing content selection and map construction based on user preferences. This range of possibilities, from adaptable to adaptive systems, allows not only the consideration of user requirements, but also spatial and temporal context. It is explained which effort is required in geo-enabling spatially relevant data to make existing spatial and non-spatial information available for location-based services.

Chapter 10  
(seventh paper)

The last paper illustrates why methods for the dynamic generation of maps for mobile devices must distinguish between 'thematic' foreground and 'topographic' background information to allow real-time access (upcoming topic 1, Table 3.2), generalisation and visualisation of dynamic information. Whilst topographic background information can be generalised in advance, dynamically requested foreground information must be generalised on-the-fly.

As explained above, an important consideration for information presentation on mobile devices is the replacement of static maps through *interactive and dynamic behaviour*. To compensate for the small screen sizes available for map presentation, user interactions such as panning and zooming can be applied to access overview as well as detailed information. Methods based on the utilisation of hierarchical data structures, which allow a rapid change between several presentation levels, were presented in paper seven. The approaches described focus on the usage of either the quadtree data structure, the derivation of a hierarchical dendrogram structure from clustering algorithms or the application of a hierarchical subdivision based on the Strahler ordering of a flow accumulation network. The last example has the further advantage of maintaining spatial relationships between foreground and background information, which potentially unlocks more valuable information.

## 11.2 Outlook and further work

Future research topics in automated generalisation will be strongly influenced by the requirements of the growing breadth of applications which apply spatial data. For example, the requirements for the representation of spatial information have been driven by new technologies such as mobile information systems, positioning techniques, service architectures, multimedia and interaction possibilities or web usage for data transfer and presentation. A *main thesis* of this work is that further research on automated generalisation must address the expectation that there will be a demand for both *traditional topographic mapping* as well as *user centred adaptive mapping*, depending on the diversity of applications and user requirements. This means user centred adaptive mapping will not simply replace topographic mapping rather, it will be a complementary application of spatial data. The characterisation of these two different scenarios will lead to the derivation of future generalisation research tasks.

Traditional topographic mapping is aimed at general, unspecified, multi-purpose usage. The maps are static trying to present a maximum number of real-world objects and their interrelations. Topographic mapping places high demands on readability and graphical quality and final products can be considered not only as maps, but as works of art. Production cycles are measured within days, months or years and users typically retain maps for decades. Automated generalisation functionality for this mapping type is therefore less constrained by time, but more by the satisfaction of cartographic constraints which preserve legibility, shape, structure and topology. Research tasks for topographic mapping are categorised in Table 3.1. Major topics include explicit modelling of horizontal and vertical relations within MRDBs, the incorporation of this enriched information through generalisation operators during initial and incremental generalisation, the automated evaluation of cartographic solutions and the accompanying need for automated process orchestration with generalisation operator selection and parameterisation.

*Traditional  
topographic mapping*

User centred adaptive mapping is aimed at specific, context and purpose dependent usage. The maps are interactive at least through zooming and panning as well as dynamic, for instance, through the consideration of location, in attempting to select and display relevant real-world objects with their important interrelations. The generation of adaptive mapping is highly dependent on purpose and user dependent content selection. Production cycles are measured in milliseconds or seconds and users exploit the maps immediately.

*User centred adaptive  
mapping*

Automated generalisation functionality for this mapping type has therefore high requirements on time and less effort need be invested in finessing cartographic outputs, because further detailed information can be reached immediately through interactions such as zooming or by exploring linked multimedia information. Research tasks for user centred adaptive mapping include modelling of context as well as the user, the design of user interfaces and presentation metaphors, the development of real-time algorithms, and the exploitation of hierarchical data structures. For example, presentation metaphors might include alternative map presentations such as perspective, ego-centric views, fish-eye zoom and liquid browsing.

Research on automated generalisation has in the past mainly dealt with methods on knowledge acquisition from cartographers and maps to reproduce the manual production and generalisation process. In user centred adaptive mapping the acquisition process should now concentrate on the user, trying to capture interactions and context information which have a direct influence on map creation and generalisation. In considering user intentions and activities, the selection and prioritisation of cartographic constraints must also become dynamic.

Furthermore, new mapping possibilities open the door for alternative generalisation methods such as the usage of extended graphical variables, including *opacity*. Less important features need not necessarily be removed, instead they may be presented with higher opacity values. These techniques become important with infinitely variable zooming and continuous generalisation.

*Cartographic  
communication*

Automated generalisation can be interpreted in an extended view as a tool to support *cartographic communication*. As Brodersen (2001, p. 35) says: "The aim of cartographic generalisation is communication with the least possible uncertainty (in the shortest possible time)." The fundamental question - what is a good map or good generalisation - can then be answered with a measure of how well it delivers information to its readers. An extension of this idea might be the development of maps which support communication between people, not only in a uni-directional way. Communication normally takes place between people, therefore it is interesting to offer suitable map frameworks, which also allow the user to communicate information. Examples can be found for instance, in user placement of POIs in Google Earth or within an application such as Open Map Sihlwald (<http://www.openmapsihlwald.ch>), which partially transfers the process of design and content presentation to the user.

Finally, I would like to refer to the keynote of Mackaness (2006) presented at the AutoCarto 2006 conference with his characterisation of current failures and visions for the future in automated

cartography. Following his argument (p. 245) that "... automated cartography is in a bush of ghosts - trapped in tradition and in a bit of a wilderness ...", I have to agree that there exist in automated cartography a mixture of unsolved problems and new challenges. As an example of unsolved problems, the "capture once, use many times" approach can be mentioned, which was and is an ultimate aim of automating traditional topographic map production. On the other hand a complete new way of thinking is necessary, as [Mackaness \(2006, p. 248\)](#) says "the introduction of technology fundamentally changes the way we go about doing a task". In this sense research on user centred adaptive maps on mobile devices provides a promising example of the transfer of research from the traditional cartographic domain to the challenging area of research driven by the rapid growth of new mapping opportunities.



## Bibliography

- Abderhalden, W. and Krug, K. (2003). Visitor monitoring in the swiss national park - towards appropriate information for the wireless consumer. In *Proceedings 10th International Conference on Information Technology and Travel & Tourism*, Helsinki, Finland. CD-ROM. 156, 160, 163
- Agarwal, P. (2005). Ontological considerations in giscience. *International Journal of Geographical Information Science*, 19(5):501–536. 160
- AGENT (1999). Report C1 – Selection of basic measures. Technical report, Institut Géographique National France, University of Edinburgh, University of Zurich and Institut National Polytechnique de Grenoble. 149
- Allouche, M. K. and Moulin, B. (2002). Une approche à base de cartes de kohonen pour la généralisation cartographique de zones de points de forte densité. In Ruas, A., editor, *Généralisation et représentation multiple*, Hermès Science, pages 337–352. 138
- Anders, K. and Sester, M. (2000). Parameter-free cluster detection in spatial databases and its application to typification. In *International Archives of Photogrammetry and Remote Sensing*, volume 33, Amsterdam, Netherlands. CD-ROM. 138, 175
- Anders, K.-H. (2003). A hierarchical graph-clustering approach to find groups of objects. In *Proceedings of the 6th Workshop on Progress in Automated Map Generalization*, Paris, France. CD-ROM. 138
- Andrienko, N. and Andrienko, G. (1999). Interactive maps for visual data exploration. *International Journal of Geographical Information Science*, 13(4):355–374. 10
- Badard, T. and Braun, A. (2003). Oxygene: An open framework for the deployment of geographic web services. In *Proceedings of the 21st International Cartographic Conference*, Durban, South Africa. CD-ROM. 60

- Badaud, J., Witkin, A. P., and Duda, R. O. (1986). Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on pattern analysis and image processing*, 8(1):183–193. [131](#)
- Bader, M. (2001). *Energy Minimizing Methods for Feature Displacement in Map Generalization*. PhD thesis, Department of Geography, University of Zurich, Switzerland. [46](#), [71](#), [122](#), [125](#)
- Balboa, J. and Lopez, F. (2000). Frequency filtering of linear features by means of wavelets: A method and an example. *Cartographic Journal*, 37(1):39–49. [131](#)
- Barnes, S. (2002). The mobile commerce value chain: Analysis and future developments. *International Journal of Information Management*, 22:91–108. [156](#)
- Barrault, M., Bader, M., and Weibel, R. (2000). Topology preserving conflict removal between symbolized roads in cartographic generalization: Extending snakes methods. In *Abstracts of GI-Science Conference*, Savannah, Georgia, USA. CD-ROM. [122](#)
- Barrault, M., Regnauld, N., Duchêne, C., Haire, K., Baeijs, C., Demazeau, Y., Hardy, P., Mackaness, W., Ruas, A., and Weibel, R. (2001). Integrating multi-agent, object-oriented and algorithmic techniques for improved automated map generalization. In *Proceedings 20th International Cartographic Conference*, Beijing, China, pages 2110–2116. [37](#), [83](#), [84](#)
- Bartram, L., Ho, A., Dill, J., and Henigman, F. (1995). The continuous zoom: A constrained fisheye technique for viewing and navigating large information spaces. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, ACM Press, C/JISTPS, Pittsburgh, Pennsylvania, pages 207–216. [31](#)
- Beard, M. K. (1991). Constraints on rule formation. In Buttenfield, B. P. and McMaster, R. B., editors, *Map Generalization: Making Rules for Knowledge Representation*, Longman, London, UK, pages 121–135. [45](#), [82](#)
- Berger, A. (1974). Ein Generalisierungsmodell zur Siedlungsdarstellung von Häuserreihen in topographischen Karten. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(65):7–17. [43](#)
- Berman, D., Bartell, J., and Salesin, D. (1994). Multiresolution painting and composition. In *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH94*, Orlando, Florida, USA, pages 85–90. [167](#)



- Bertolotto, M. and Egenhofer, M. (1999). Progressive vector transmission. In *Transactions of the ACMGIS99*, Kansas City, Missouri, USA, pages 152–157. 29
- Bertolotto, M. and Egenhofer, M. (2001). Progressive transmission of vector map data over the world wide web. *GeoInformatica*, 5(4):345–373. 29
- Board, C. (1967). Maps as models. In Chorley, R. J. and Haggett, P., editors, *Models in Geography*, Methuen & Co, London, UK, pages 671–725. 25
- Board on Earth Sciences and Resources (BESR) (2006). *Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum*. The National Academies Press. 12, 13
- Bobrich, J. (1996). *Ein neuer Ansatz zur kartographischen Verdrängung auf der Grundlage eines mechanischen Federmodells*. PhD thesis, Deutsche Geodätische Kommission, München, Reihe C, H. 455. 46, 122
- Bobrich, J. (2001). Cartographic map generalization in urban districts. In *Proceedings of the GIS Research UK, 9th Annual Conference*, Glamorgan, Wales, pages 513–515. 44
- Bobzien, M., Burghardt, D., and Petzold, I. (2005a). Automatische Ableitung digitaler Vektormodelle - Projekt DRIVE. *Mitteilungen des Bundesamtes für Kartographie und Geodäsie: Arbeitsgruppe Automation in der Kartographie, Tagung 2004*, 34:9–20. 55
- Bobzien, M., Burghardt, D., and Petzold, I. (2005b). Re-Generalisation and Construction - Two Alternative Approaches to Automated Incremental Updating in MRDB. In *Proceedings of 22nd International Cartographic Conference*, La Coruña, Spain. CD-ROM. 55, 108, 113, 185
- Bobzien, M., Burghardt, D., Petzold, I., Neun, M., and Weibel, R. (2006a). Multi-representation databases with explicitly modelled intra-resolution, inter-resolution and update relations. In *Proceedings of AutoCarto*, Vancouver, Washington, USA. CD-ROM. 41
- Bobzien, M., Burghardt, D., Petzold, I., Neun, M., and Weibel, R. (accepted 2007). Multi-representation databases with explicitly modelled horizontal, vertical and update relations. *Cartography and Geographic Information Science*. 54
- Bobzien, M., Burghardt, D., Petzold, I., and Weibel, R. (2006b). Ableitung von digitalen Vektormodellen Ergebnisse des Projektes DRIVE. *Kartographische Nachrichten*, 56(5):254–262. 55

- Boffet, A. and Rocca Serra, S. (2001). Identification of spatial structures within urban blocks for town characterization. In *Proceedings 20th International Cartographic Conference*, Beijing, China, pages 1974–1983. 88
- Boyle, A. R. (1970). The quantised line. *The Cartographic Journal*, 7(2):91–94. 43
- Brassel, K. and Weibel, R. (1988). A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Systems*, 2(3):229–244. xii, 5, 36, 43, 44
- Brassel, K. E. (1985). Strategies and data models for computer-aided generalization. *International Yearbook of Cartography*, 25:11–29. 43
- Brenner, C. and Sester, M. (2003). Continuous generalization for small mobile displays. In *International Workshop on Next Generation Geospatial Information*, Cambridge (Boston), Massachusetts, USA. 27
- Brenner, C. and Sester, M. (2005). Continuous generalization for small mobile displays. In Agouris, P. and Croituru, A., editors, *Next Generation Geospatial Information*, ISPRS Book Series, Vol 1, No. 2, Taylor & Francis Group, London, UK, pages 33–42. 27
- Brodersen, L. (2001). Maps as communication. theory and methodology in cartography. Technical report, Technical Report No.14. National Survey and Cadastre Denmark. 188
- Burger, W. and Burge, M. J. (2005). *Digitale Bildverarbeitung*. Springer, Berlin, Heidelberg, New York. 29
- Burghardt, D. (2000). *Automatisierung der kartographischen Verdrängung mittels Energieminimierung*. PhD thesis, Department of Geodasy, Technical University of Dresden, Germany. 122, 125
- Burghardt, D. (2005). Controlled line smoothing by snakes. *GeoInformatica*, 9(3):237–252. 55
- Burghardt, D., Bobzien, M., Petzold, I., and Weibel, R. (2005a). Cartographic generalisation of large scale maps with axpand. In *Proceedings of International Symposium on Generalization of Information*, Berlin, Germany, pages 147–159. 41, 55
- Burghardt, D. and Cecconi, A. (2003). Mesh simplification for building selection. In *Proceedings 6th ICA Workshop on Progress in Automated Map Generalization*, Paris, France. CD-ROM. 28

- Burghardt, D. and Cecconi, A. (2007). Mesh simplification for building typification. *International Journal of Geographical Information Science*, 21(3):283–298. 55
- Burghardt, D., Edwardes, A., and Mannes, J. (2004a). An architecture for automatic generalisation of mobile maps. In *Proceedings of 2nd Symposium on Location Based Services and TeleCartography*, Vienna, Austria, pages 33–38. 22, 38, 54
- Burghardt, D., Edwardes, A., Purves, R., and Weibel, R. (2005b). Automatische Generalisierung dynamisch generierter Karten für mobile Endgeräte. *Kartographische Nachrichten*, 55(3):119–125. 56, 57
- Burghardt, D., Edwardes, A., and Weibel, R. (2003). WebPark - Entwicklung eines mobilen Informationssystems für Erholungs- und Naturschutzgebiete. *Kartographische Nachrichten*, 53(2):58–64. 172, 181
- Burghardt, D. and Mathur, A. (2004). Derivation of digital vector models project drive. In *Proceedings 7th ICA Workshop on Generalisation and Multiple Representation*, Leicester, UK. CD-ROM. 55
- Burghardt, D. and Meier, S. (1997). Cartographic displacement using the snakes concept. In W. Förstner, W. and Plümer, L., editors, *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*, Birkhäuser Verlag, Basel, Switzerland, pages 59–71. 46, 71, 122, 124
- Burghardt, D. and Neun, M. (2006). Automated sequencing of generalisation services based on collaborative filtering. In Raubal, M., Miller, H. J., Frank, A. U., and Goodchild, M., editors, *4th International Conference, Geographic Information Science*, IfGIprints 28, pages 41–46. 55, 183
- Burghardt, D., Neun, M., and Weibel, R. (2005c). Generalization services on the web - a classification and an initial prototype implementation. *Cartography and Geographic Information Science*, 32(4):257–268. 55
- Burghardt, D., Purves, R., and Edwardes, A. (2004b). Techniques for on-the-fly generalisation of thematic point data using hierarchical data structures. In *Proceedings of the GIS Research UK 12th Annual Conference, Norwich, UK.*, pages 60–63. 22, 57, 138, 179
- Burghardt, D., Schmid, S., and Stoter, J. (2007). Investigations on cartographic constraint formalisation. In *Proceedings 10th*

- ICA Workshop on Generalisation and Multiple Representation*, Moscow, Russia. CD-ROM. 53
- Burghardt, D. and Steiniger, S. (2005). Usage of principal component analysis in the process of automated generalisation. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. 117
- Buttenfield, B. (2002). Transmitting vector geospatial data across the internet. In Egenhofer, M. and Mark, D., editors, *Proceedings GIScience, Lecture Notes in Computer Science*, volume 2748, pages 51–64. 29
- Buttenfield, B. P. and McMaster, R. B., editors (1991). *Map Generalization. Making Rules for Knowledge Representation*. Longman, London, UK. 43
- Buttenfield, B. P., Weber, C. R., and Leitner, M. (1991). How does a cartographic object behave? Computer inventory of topographic maps. In *Proceedings GIS/LIS'91*, Atlanta, USA, pages 891–900. 44, 51
- Cartwright, W., Miller, S., and Pettit, C. (2004). Geographical visualization: Past, present and future development. *Journal of Spatial Science*, 49(1):25 – 36. 10
- Cartwright, W. E., Peterson, M. P., and Gartner, G. (1999). *Multi-media Cartography*. Heidelberg: Springer-Verlag. 10
- Cecconi, A. (2003). *Improved Automated Generalization for On-Demand Web Mapping by Multiscale Databases*. PhD thesis, Department of Geography, University of Zurich. xi, 28, 54
- Cecconi, A., Burghardt, D., and Weibel, R. (2005). Integration von Multirepräsentationsdatenbanken in den Generalisierungsprozess für die Internetkartographie. *Kartographische Nachrichten*, 55(1):11–17. 54
- Cecconi, A. and Galanda, M. (2002). Adaptive zooming in web cartography. *Computer Graphics Forum*, 21(4):787–799. 27, 160, 172
- Chen, Z., Lin, F., Liu, H., Liu, Y., Ma, W.-Y., and Wenying, L. (2002). User intention modeling in web applications using data mining. In *World Wide Web: Internet and Web Information Systems*, Vol. 5, Kluwer Academic Publishers, pages 181–191. 18, 19
- Cheverst, K., Davies, N., Mitchell, K., and Friday, A. (2000). Experiences of developing and deploying a context-aware tourist

- guide: the guide project. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, ACM Press, New York, NY, USA, pages 20–31. [164](#), [169](#)
- Christ, F. (1977). Vollautomatische und halbautomatische Generalisierung, Symbolisierung und Lichtzeichnung einer topographischen Karte kleinen Maßstabs. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(71):27–46. [43](#)
- Christ, F. (1979). Ein Programm zur vollautomatischen Verdrängung von Punkt- und Linienobjekten bei der kartographischen Generalisierung. *Internationales Jahrbuch für Kartographie XIX*, pages 41–64. [43](#)
- Christophe, S. and Ruas, A. (2002). Detecting building alignments for generalisation purposes. In Richardson, D. and van Oosterom, P., editors, *Advances in Spatial Data Handling. 10th International Symposium on Spatial Data Handling*, Springer, Berlin, Heidelberg, pages 419–432. [87](#), [107](#), [153](#)
- Clarke, K., Cippolletti, P., and Olsen, G. (1993). Empirical comparison of two line enhancement methods. In *Proceedings 11 International Symposium on Computer-Assisted Cartography, Auto-Carto*, Minneapolis, USA, pages 72–81. [131](#)
- Clarke, K. C. (2004). Mobile mapping and geographic information systems. *Cartography and Geographic Information Science*, 31(3):131–136. [16](#), [21](#)
- Cooper, A. (2003). The concepts of incremental updating and versioning. In *Proceedings of the 21st International Cartographic Conference*, Durban, South Africa, pages 855–857. [111](#)
- Coors, V., Elting, C., Kray, C., and Laakso, K. (2005). Presenting route instructions on mobile devices - from textual directions to 3d visualization. In Dykes, J., MacEachren, A. M., and Kraak, M.-J., editors, *Exploring Geovisualization*, Pergamon, pages 529–550. [33](#)
- Cromley, R. G. and Campbell, G. M. (1992). Noninferior bandwidth line simplification: Algorithm and structural analysis. *Geographical Analysis*, 23(1):25–38. [45](#)
- Cuthbert, A. (1998). User interaction with geospatial data. Technical report, In OpenGIS Project Document 98-060. [xii](#), [5](#), [35](#)
- Das, S., Chong, E. I., Eadon, G., and Srinivasan, J. (2004). Supporting Ontology-based Semantic Matching in RDBMS. In *Proceedings of the 30th International Conference on Very Large Data Bases*, Toronto, Canada, pages 1054–1065. [183](#)

- de Berg, M., Bose, P., Cheong, O., and Morin, P. (2004). On simplifying dot maps. *Computational Geometry*, 27:43–62. [138](#), [167](#)
- De Floriani, L., Magillo, P., and Puppo, E. (2000). VARIANT: A system for terrain modeling at variable resolution. *GeoInformatica*, 4(3):287–315. [29](#)
- De Lucia, A. A. and Black, R. T. (1987). A comprehensive approach to automatic feature generalization. In *Proceedings 13th International Cartographic Conference*, Morelia, Mexico, pages 173–191. [43](#), [44](#)
- Devogele, T., Trevisan, J., and Raynal, L. (1996). Building a multi-scale database with scale-transition database. In Kraak, M. and Molenaar, M., editors, *Advances in GIS research II*, Taylor & Francis, London, UK, pages 559–570. [102](#)
- Dey, A. (2001). Understanding and using context. *Personal and Ubiquitous Computing*, 5(1):4–7. [18](#)
- Dias, E., Beinart, E., and Scholten, H. (2004). Effects of mobile information sharing in natural parks. In *Proceedings of 18th International Conference Informatics for Environmental Protection, ENVIROINFO*, Tricorne Editions, Geneva, Switzerland, pages 11–25. [156](#), [163](#)
- Dias, E. and Edwardes, A. J. (2005). Information Flows in Nature Areas - Information Needs and Data supply for Location-based Services in Nature Areas. In *Proceedings GIS-Planet*, Lisbon, Portugal. CD-ROM. [160](#)
- DiBiase, D., MacEachren, A. M., Krygier, J. B., and Reeves, C. (1992). Animation and the role of map design in scientific visualization. *Cartography and Geographic Information Systems*, 19(4):201–214. [10](#)
- Dodson, S. (2005). Get mapping - as mapmaking becomes big business, citizen cartographers are creating free personal alternatives. *Guardian, online newspaper*. April 7, 2005. [9](#)
- Donelson, W. C. (1978). Spatial management of information. In *Proceedings of the 5th annual conference on Computer graphics and interactive techniques, SIGGRAPH78*, ACM Press, Atlanta, Georgia, USA, pages 203–209. [31](#)
- Dorling, D. (1992). Stretching space and splicing time: from cartographic animation to interactive visualization. *Cartography and Geographic Information Systems*, 19(4):215–227. [10](#), [30](#)



- Dorling, D. (1996). *Area Cartograms: Their Use and Creation*. Concepts and Techniques in Modern Geography (CATMOG) No. 59, University of East Anglia, Norwich. 30
- Dougenik, J. A., Chrisman, N. R., and Niemeyer, D. R. (1985). An algorithm to construct continuous area cartograms. *The Professional Geographer*, 47(1):75–81. 30
- Douglas, D. H. and Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its character. *The Canadian Cartographer*, 10(2):112–123. 43
- Dourish, P. (2004). What we talk about when we talk about context. *Personal Ubiquitous Computing*, 8(1):19–30. 162
- Doyle, J., Weakliam, J., Bertolotto, M., and Wilson, D. (2006). A multimodal interface for personalising spatial data in mobile gis. In Manolopoulos, Y., Filipe, J., Constantopoulos, P., and Cordeiro, J., editors, *Proceedings of the 8th International Conference on Enterprise Information Systems: Databases and Information Systems Integration*, Paphos, Cyprus, pages 71–78. 18, 19
- Dürst, M. J. (1990). *A New Method for Image Compression and Progressive Transmission*. PhD thesis, Doctoral Dissertation, Faculty of Science, University of Tokyo, Japan. 28
- Dürst, M. J. and Kunii, T. L. (1991). Progressive transmission increasing both spatial and gray scale resolution. In *Proceeding of International Conference on Multimedia Information Systems*, Singapore, pages 175–186. 29
- Duchêne, C. (2003). Automated map generalisation using communicating agents. In *Proceedings of 21st International Cartographic Conference*, Durban, South Africa, pages 160–169. 46
- Duchêne, C. (2004). The CartACom model : a generalisation model for taking relational constraints into account. In *Proceedings 7th ICA Workshop on Generalisation and Multiple Representation*, Leicester, UK. CD-ROM. xii, 38, 46, 103
- Duchêne, C. (2001). Road generalisation using agents. In *Proceedings of the GIS Research UK, 9th Annual Conference*, Glamorgan, Wales, pages 325–328. 135
- Duda, R. and Hart, P. (1973). *Pattern Classification and Scene Analysis*. John Wiley, New York. 43
- Duda, R., Hart, P., and Stork, D. (2000). *Pattern Classification*. John Wiley, New York, 2nd edition. 92



- Dutton, G. (1996). Encoding and handling geospatial data with hierarchical triangular meshes. In Kraak, M. J. and Molenaar, M., editors, *Advances in GIS Research II*, Taylor & Francis, London, pages 505–518. 177
- Dutton, G. (1998). *A hierarchical coordinate system for geoprocessing and cartography*. Lecture Notes in Earth Science 78, Springer-Verlag, Berlin. 45
- Dykes, J., MacEachren, A. M., and Kraak, M.-J. (2005). *Exploring geovisualization*. Amsterdam, Elsevier. 10
- Edwardes, A. (2007). *Re-placing Location: Geographic Perspectives in Location Based Services*. PhD thesis, Department of Geography, University Zurich, Switzerland. xi, 21, 32, 33, 57, 62
- Edwardes, A., Burghardt, D., Bobzien, M., Harrie, L., Lehto, L., Reichenbacher, T., Sester, M., and Weibel, R. (2003a). Map generalisation technology: Addressing the need for a common research platform. In *Proceedings of 21st International Cartographic Conference*, Durban, South Africa, pages 170–179. 22, 40, 54, 60
- Edwardes, A., Burghardt, D., Dias, E., Purves, R. S., and Weibel, R. (2005a). Geo-enabling spatially relevant data for mobile information use and visualisation. In *Lecture Notes in Computer Science. 5th International Workshop on Web and Wireless Geographical Information Systems*, Springer-Verlag, pages 78–92. 56, 57
- Edwardes, A., Burghardt, D., and Neun, M. (2005b). Interoperability in map generalisation research. In *Proceedings of International Symposium on Generalization of Information, ISGI*, Berlin, pages 160–174. 55
- Edwardes, A., Burghardt, D., and Neun, M. (2007). Experiments to build an open generalisation system. In Ruas, A., Mackaness, W. A., and Sarjakoski, T., editors, *Generalisation of Geographic Information: Cartographic Modelling and Applications*, Series of International Cartographic Association, Elsevier Science, pages 161–175. 22, 36, 38, 39, 40, 55
- Edwardes, A., Burghardt, D., and Weibel, R. (2003b). WebPark - Location Based Services for species search in recreation area. In *Proceedings of 21st International Cartographic Conference*, Durban, South Africa, pages 1012–1022. xii, 20, 21, 34, 35, 57, 160
- Edwardes, A., Burghardt, D., and Weibel, R. (2005c). Portrayal and generalisation of point maps for mobile information services. In

- Meng, L., Zipf, A., and Reichenbacher, T., editors, *Map-based Mobile Services*, Springer Verlag, pages 11–30. [xi](#), [21](#), [24](#), [25](#), [57](#)
- Egenhofer, M. and Herring, J. (1991). Categorizing binary topological relationships between regions, lines, and points in geographic databases. Technical report, University of Maine, Orono. [106](#)
- Egenhofer, M.-J., Clementini, E., and Felice, P.-D. (1994). Evaluating inconsistencies among multiple representations. In *Proceedings of 6th International Symposium on Spatial Data Handling*, Vol. 2, Edinburgh, Scotland, pages 901–920. [45](#)
- Egenhofer, M. J. and Franzosa, R. (1991). Point-set spatial topological relations. *International Journal of Geographical Information Systems*, 5(2):161–174. [106](#)
- Estivill-Castro, V. and Lee, I. (2000). AMOEBA: Hierarchical clustering based on spatial proximity using delaunay diagram. In *Proceedings of the 9th International Symposium on Spatial Data Handling*, Beijing, China, pages 7A.26–7A.41. [138](#)
- Feldmann, H.-U. and Kreiter, N. (2006). Neuaufbau der schweizerischen Landeskarte: Inhalt und Kartengrafik. *Kartographische Nachrichten*, 3:115–121. [xi](#), [14](#), [15](#)
- Filli, F., Haller, R., Moritzi, M., Negri, M., Obrecht, J.-M., Robin, K., and Schuster, R. (2000). Die Singvögel im Schweizerischen Nationalpark: Verbreitung anhand GIS-gestützter Habitatmodell. *Jahresbericht der Naturforschenden Gesellschaft Graubünden, Chur*, 109:47–90. [163](#)
- Fischer, G. (2001). User modeling in human-computer interaction. *Contribution to the 10th Anniversary Issue of the Journal "User Modeling and User-Adapted Interaction (UMUAI)"*, 11(1-2):65–86. [18](#)
- Fitzke, J., Greve, K., Müller, M., and Poth, A. (2004). Building SDIs with free software – the degree project. In *Proceedings of 7th Conference on Global Spatial Data Infrastructure*, Bangalore, India, pages 1–13. [61](#)
- Fornefeld, M., Oefinger, P., and Rausch, U. (2003). The market for geospatial information: Potentials for employment, innovation and value added. Technical report, MICUS Management Consulting GmbH. [xi](#), [13](#)
- Frank, A. U. and Timpf, S. (1994). Multiple representations for cartographic objects in a multiscale tree - an intelligent graphical zoom. *Computers and Graphics*, 18(6):823–829. [45](#)

- Fritsch, E. (1997). Utilisation de la courbure pour la généralisation du linéaire routier. *Bulletin d'Information*, n. 66. IGN, France. 131
- Fritsch, E. and Lagrange, J. P. (1995). Spectral representation of linear features for generalisation. In *Proceedings of COSIT95*, Austria, Springer Verlag, pages 157–171. 131
- Furnas, G. (1981). The fisheye view: new look at structured files. Technical report, Bell Laboratories technical memorandum, #81-11221-9. 31
- Furnas, G. (1986). Generalized fisheye views. In *Human Factors in computing systems, CHI '86 conference proceedings*, ACM, New York, pages 16–23. 31
- Gaffuri, J. and Trévisan, J. (2004). Role of urban patterns for building generalisation: An application of agent. In *Proceedings of 7th ICA Workshop on Generalisation and Multiple Representation*, Leicester, Great Britain. CD-ROM. 91
- Gahegan, M. (2005). Beyond tools: visual support for the entire process of giscience. In Dykes, J., MacEachren, A. M., and Kraak, M. J., editors, *Exploring GeoVisualization*, Elsevier, Amsterdam, The Netherlands, pages 83–99. 12
- Galanda, M. (2003). Modelling constraints for polygon generalization. In *Proceedings 6th ICA Workshop on Progress in Automated Map Generalization*, Paris, France. CD-ROM. 54
- Galanda, M. and Weibel, R. (2003). Using an energy minimization technique for polygon generalization. *Cartography and Geographic Information Science*, 30(3):259–275. 46, 71
- Gastner, M. T. and Newman, E. J. (2004). Diffusion-based method for producing density equalizing maps. In *Proceedings of the National Academy of Sciences of the United States of America*, pages 7499–7504. 30
- Gastner, M. T. and Newman, E. J. (2005). Density-equalizing map projections: Diffusion-based algorithm and applications. In *Proceedings of the 8th International Conference on Geocomputation*, Ann Arbor, Michigan, USA. CD-ROM. 30
- GKG-KOGIS (2003). Umsetzungskonzept zur Strategie für Geoinformation beim Bund. Technical report, Koordination der Geoinformation und der geographischen Informationssysteme (KOGIS), Switzerland. 14

- Gold, C. M. and Thibault, D. (2001). Map generalization by skeleton retraction. In *Proceedings International Cartographic Association*, Beijing, China, pages 2072–2081. 45
- Goodchild, M. (1992). Geographical information science. *International Journal of Geographical Information Systems*, 6(1):31–45. 12
- Goodchild, M. F. (2005). GIS and modeling overview. In Maguire, D. J., Batty, M., and Goodchild, M. F., editors, *GIS, Spatial Analysis, and Modeling*, Redlands, CA: ESRI Press, pages 1–18. 42
- Gottschalk, H.-J. (1974). Automatische Generalisierung von Siedlungen, Verkehrswegen, Höhenlinien, Wasserläufen und Vegetationsgrenzen für eine kleinmaßstäbige topographische Karte. *Zeitschrift für Vermessungswesen*, 99:338–342. 43
- Gottschalk, H.-J. (1973). Die automatische Ableitung einer Siedlungsdarstellung im Maßstab 1:200'000 aus im Maßstab 1:50'000 digitalisierten Flächen und Einzelhäusern. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(71):19–29. 43
- Grünreich, D. (1993). Generalization in GIS environment. In *Proceedings of the 16th International Cartographic Conference*, Cologne, Germany, pages 203–210. 44
- Guha, S., Rastogi, R., and Shim, K. (1999). ROCK: A robust clustering algorithm for categorical attributes. In *Proceedings of the International Conference on Data Engineering*, Sydney, Australia, pages 512–521. 138
- Gusein-Zade, S. M. and Tikunov, V. (1993). A new technique for constructing continuous cartograms. *Cartography and Geographic Information Systems*, 20(3):167–173. 30
- Hake, G. (1985). *Kartographie II - Thematische Karten, Atlanten, kartenverwandte Darstellungen, Kartenredaktion und Kartentechnik, rechnergestützte Kartenherstellung, Kartenauswertung, Kartengeschichte*. Walter de Gruyter, Berlin, New York. 43
- Hake, G., Grünreich, D., and Meng, L. (2002). *Kartographie*. Walter de Gruyter, 8. Auflage., Berlin, New York, 8th edition. 51
- Hake, G. and Hoffmeister, E.-D. (1978). Computerunterstützte Gebäudegeneralisierung. *Kartographische Nachrichten*, 28:47–55. 43

- Hampe, M., Anders, K.-H., and Sester, M. (2003). MRDB applications for data revision and real-time generalisation. In *Proceedings of the 21st International Cartographic Conference*, Durban, South Africa, pages 192–202. 102
- Hardy, P. and Lee, D. (2005). GIS-based generalization and multiple representation of spatial data. In *Proceedings International Symposium on Generalization of Information*, ISGI, Berlin, Germany, pages 175–190. xii, 36, 37
- Harrie, L. (1999). The constraint method for solving spatial conflicts in cartographic generalization. *Cartography and Geographic Information Science*, 26(1):55–69. 71, 83
- Harrie, L. (2001). *An Optimisation Approach to Cartographic Generalisation*. PhD thesis, Department of Technology and Society, Lund University. 46, 123
- Harrie, L. and Hellström, A. K. (1999). A prototype system for propagating updates between cartographic data sets. *The Cartographic Journal*, 36(2):133–140. 109
- Harrie, L. and Johansson, M. (2003). Real-time data generalisation and integration using java. *Geoforum Perspektiv*, 2:29–34. 38
- Harrie, L. and Sarjakoski, T. (2000). Generalisation of vector data sets by simultaneous least squares adjustment. In *International Archives of Photogrammetry and Remote Sensing*, Volume XXXIII, Part A4, Amsterdam, The Netherlands, pages 340–347. 123
- Harrie, L. and Sarjakoski, T. (2002). Simultaneous graphic generalisation of vector data sets. *GeoInformatica*, 6(2):233–261. 31
- Harrie, L., Sarjakoski, T., and Lehto, L. (2002). A variable-scale map for small-display cartography. In *Proceedings of the Joint International Symposium on Geospatial Theory, Processing and Applications*, Ottawa, Canada. CD-ROM. 172
- Harrie, L. and Weibel, R. (2007). Modelling the overall process of generalisation. In Ruas, A., Mackaness, W. A., and Sarjakoski, T., editors, *Generalisation of Geographic Information: Cartographic Modelling and Applications*, Series of International Cartographic Association, Elsevier Science, pages 67–87. 45
- Haunert, J.-H., Anders, K.-H., and Sester, M. (2006). Hierarchical structures for rule based incremental generalisation. In *Joint ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data*, Hannover, Germany. CD-ROM. 109

- Haunert, J.-H. and Sester, M. (2005). Propagating updates between linked datasets of different scales. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. 109
- Hearnshaw, H. and Unwin, D. (1994). *Visualization in Geographical Information Systems*. John Wiley & Sons, New York. 12
- Heinzle, F., Kopczynski, M., and Sester, M. (2003). Spatial data interpretation for the intelligent access to spatial information in the internet. In *Proceedings of 21st International Cartographic Conference*, Durban, South Africa, pages 1040–1050. 91
- Heinzle, F., Sester, M., and Anders, K. (2005). Graph based approach for recognition of patterns and implicit information in road networks. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. 91, 107
- Herzog, A. (2003). Developing cartographic applets for the internet. In Peterson, M. P., editor, *Maps and the Internet*, pages 115–128. 30
- Hoffmeister, E.-D. (1978). Programmgesteuerte Generalisierung für die Topographische Karte 1:25 000. *Nachrichten aus dem Karten- und Vermessungswesen, Frankfurt a. M.*, 1(75). 44
- Højholt, P. (1995). Generalization of built-up areas using kohonen networks. In *Proceedings of Eurocarto XIII*, Ispra, Italia, pages 177–182. 138
- Højholt, P. (1998). Solving local and global space conflicts in map generalisation using finite element method adapted from structural mechanics. In *Proceedings of the 8th International Symposium on Spatial Data Handling*, Vancouver, Canada, pages 679–689. 122
- Højholt, P. (2000). Solving space conflicts in map generalization: Using a finite element method. *Cartography and Geographic Information Science*, 27(1):65–74. 46, 71
- Hoppe, H. (1996). Progressive meshes. *Computer Graphics, (Annual Conference Series)*, 30:99–108. 28, 141
- Hoppe, H. (1998). Smooth view-dependent level-of-detail control and its application to terrain rendering. *IEEE Visualization*, pages 35–42. 28, 29, 30
- Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., and Stuetzle, W. (1993). Mesh Optimization. *Computer Graphics, (Annual Conference Series)*, 27:19–26. xv, 139, 140



- Horton, R. E. (1945). Erosional development of streams and their drainage basins - hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin*, 56:275–370. [43](#)
- Hotelling, H. (1931). A generalization of students ratio. *Annals of Mathematical Statistics*, 2:360–378. [98](#)
- House, D. C. and Kocmoud, C. J. (1998). Continuous cartogram construction. In *Proceedings of the Conference on Visualization*, North Carolina, USA, pages 197–204. [30](#)
- Hu, H. and Lee, D. (2004). Semantic location modeling for location navigation in mobile environment. In *Proceedings of the IEEE International Conference on Mobile Data Management*, Berkeley, California, USA, pages 52–61. [161](#)
- ICA (1973). *Multilingual Dictionary of Technical Terms in Cartography*. Franz Steiner Verlag, Wiesbaden. [24](#)
- Illert, A. and Afflerbach, S. (2004). Global schema specification. Technical report, GiMoDig-project, IST-2000-30090, Deliverable D5.3.1, Public EC report. [63](#)
- Jackson, J. E. (1991). *A users guide to principal components*. New York: John Wiley & Sons. [98](#)
- Jahard, Y., Lemarié, C., and Lecordix, F. (2003). The implementation of new technology to automate map generalisation and incremental updating processes. In *Proceedings of the 21st International Cartographic Conference*, Durban, South Africa, pages 1449–1459. [111](#)
- Jankowski, P. and Nyerges, T. (1989). Design considerations for mapkbs - a map projection knowledge-based system. *The American Cartographer*, 16(2):85–95. [44](#)
- Johannsen, T. (1977). Automatische Verfahren mit interaktivem Editing, Randanpassung, Kreuzungsbereinigung und Schriftplatzierung für eine kleinmaßstäbige karte. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(71):47–60. [43](#)
- Jones, C., Kidner, D., Luo, L., Bundy, G., and Ware, J. (1996). Database design for a multiscale spatial information system. *International Journal of Geographical Information Systems*, 10(8):901–920. [45](#), [102](#)
- Jones, C. and Ware, J. (1998). Proximity search with a triangulated spatial model. *The Computer Journal*, 41:71–83. [138](#)



- Jones, C. B., Geraint, L. B., and Ware, J. M. (1995). Map generalization with a triangulated data structure. *Cartography and Geographic Information Systems*, 22(4):317–331. 45, 138
- Jones, C. B. and Ware, J. M. (2005). Map generalization in the Web age. *International Journal of Geographical Information Science*, 19(8-9):859–870. 27, 34
- Jose, R., Moreira, A., Rodrigues, H., and Davies, N. (2003). The AROUND architecture for dynamic location-based services. *Mobile Networks and Applications*, 8(4):377–387. 161
- Kang, I., Kim, T., and Li, K. (1997). A spatial data mining method by delaunay triangulation. In *Proceedings of the 5th International Workshop on Advances in Geographic Information Systems (GIS-97)*, Las Vegas, NV, USA, pages 35–39. 138
- Karypis, G., Han, E., and Kumar, V. (1999). CHAMELEON: A hierarchical clustering algorithm using dynamic modeling. *IEEE Computer: Special Issue on Data Analysis and Mining*, 32:68–75. 138
- Kass, M., Witkin, A., and Terzopoulos, D. (1987). Snakes: Active contour models. In *Proceedings First International Conference on Computer Vision.*, London, England, IEEE Computer Soc. Press, pages 259–268. 30, 124
- Keim, D. A., Panse, C., and North, S. C. (2005). Medial-axis-based cartograms. *IEEE Computer Graphics and Applications*, 25(3):60–68. 30
- Kelly, D. and Teevan, J. (2003). Implicit feedback for inferring user preference: A bibliography. *Special Interest Group on Information Retrieval (SIGIR) Forum*, 37(2):18–28. 18
- Keyes, L. and Winstanley, A. (2001). Using moment invariants for classifying shapes on large-scale maps. *Computers, Environment and Urban Systems*, 25:119–130. 91
- Kilpeläinen, T. and Sarjakoski, T. (1995). Incremental generalization for multiple representations of geographical objects. In Müller, J. C., Lagrange, J. P., and Weibel, R., editors, *GIS and Generalization, GISDATA 1*, Taylor & Francis, pages 209–218. 109
- Kraak, M.-J. (2005). Cartography and geo-information science: an integrated approach. Technical report, 8th United Nations Regional Cartographic Conference for the Americas, New York. 12
- Kraak, M.-J. (2006). Maps, influence and being influenced. *Kartographische Nachrichten*, 3:127–131. 10, 11, 15

- Kraak, M.-J. and Brown, A. (2001). *Web Cartography: developments and prospects*. Taylor & Francis. [xi](#), [10](#)
- Kraak, M.-J., Edsall, R., and MacEachren, A. (1997). Cartographic animation and legends for temporal maps: Exploration and or interaction. In *Proceedings of the 18th International Cartographic Conference*, Stockholm, Sweden, pages 253–262. [10](#)
- Kraak, M.-J. and MacEachren, A. M. (2005). Geovisualization and GIScience. *Cartography and Geographic Information Science*, 32(2):67–68. [11](#)
- Kramer, J. (2007). Is abstraction the key to computing? *Communications of the Association for Computing Machinery (ACM)*, 50(4):36–42. [23](#)
- Kriz, K. (2001). Kartographische Ansichten im neuen Millennium. In Buzin, R. and Wintges, T., editors, *Kartographie 2001 - multidisziplinär und multidimensional. Beiträge zum 50. Deutschen Kartographentag*, Wichmann Verlag, Heidelberg, pages 228–238. [xi](#), [15](#), [16](#)
- Kuehne, S., Forbus, K., Gentner, D., and Quinn, B. (2000). SEQL: Category learning as progressive abstraction using structure mapping. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Philadelphia, PA, USA, pages 770–775. [23](#)
- Lang, T. (1969). Rules for robot draughtsmen. *The Geographical Magazine*, 42(1):50–51. [43](#)
- Lee, I. and Estivill-Castro, V. (2002). Polygonization of point clusters through cluster boundary extraction for geographical data mining. In *Proceedings of the 10th International Symposium on Spatial Data Handling*, Ottawa, Canada, pages 27–40. [138](#)
- Lehto, L. (2003). Architecture specification. GiMoDig - geospatial info-mobility service by real-time data-integration and generalisation. ist-2000-30090. Technical report, Public Deliverable D4.4.1: Final system architecture, 41 pages. [xii](#), [39](#)
- Lehto, L. and Kilpeläinen, T. (2000). Real-time generalisation of geodata in the web. In *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B4, Amsterdam, The Netherlands, pages 559–566. [38](#)
- Lehto, L. and Kilpeläinen, T. (2001). Generalizing XML-encoded spatial data on the web. In *Proceedings of 20th International Cartographic Conference*, Beijing, China, pages 2390–2396. [38](#)

- Lehto, L. and Sarjakoski, L. T. (2005). Real-time generalization of XML-encoded spatial data for web and mobile devices. *International Journal of Geographical Information Science*, 19(8-9):957–973. 22, 38
- Lehto, L. and Sarjakoski, T. (2004). An open service architecture for mobile cartographic applications. In *Proceedings of 2nd Symposium on Location Based Services and TeleCartography*, Vienna, Austria, pages 141–145. 63
- Leitner, M. and Buttenfield, B. P. (1995). Acquisition of procedural cartographic knowledge by reverse engineering. *Cartography and Geographic Information Systems*, 22(3):232–241. 44
- Lemmens, R., Wytzisk, A., de By, R., Granell, C., Gould, M., and van Oosterom, P. (2006). Integrating semantic and syntactic descriptions to chain geographic services. *IEEE Internet Computing*, 9-10:42–52. 42
- Li, Z., Yan, H., Ai, T., and CHEN, J. (2004). Automated building generalization based on urban morphology and gestalt theory. *International Journal of Geographical Information Science*, 18(5):513–534. 87, 104, 107
- Lichtner, W. (1977). EDV-unterstützte Durchführung von Verdrängungsprozessen bei der kartographischen Generalisierung in topographischen Karten. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(72):65–76. 43
- Lichtner, W. (1979). Computer-assisted processes of cartographic generalization in topographic maps. *Geo-Processing*, 1:183–199. 43
- Longley, P., Goodchild, M., Maguire, D., and Rhind, D. (2005). *Geographic Information Systems and Science*. New York: Wiley, Second Edition. 11
- Lüscher, P. and Burghardt, D. (2005). Matching von Straßendaten stark unterschiedlicher Maßstäbe und Aufbau einer MRDB. *Mitteilungen des Bundesamtes für Kartographie und Geodäsie: Arbeitsgruppe Automation in der Kartographie.*, 36:79–88. 15
- MacEachren, A. (1991). The role of maps in spatial knowledge acquisition. *Cartographic Journal*, 28(2):152–162. 11
- MacEachren, A. (1994a). *Some Truth with Maps: A primer on symbolization and design*. Association of American Geographers, Washington D. C. 160

- MacEachren, A. M. (1994b). Visualization in modern cartography: Setting the agenda. In MacEachren, A. and Taylor, D. R. F., editors, *Visualization in Modern Cartography*, Elsevier, New York, USA, pages 1–12. [11](#)
- Mackaness, W. and Beard, M. (1993). Use of graph theory to support map generalization. *Cartography and Geographic Information Systems*, 20(4):210 – 221. [45](#)
- Mackaness, W., Fisher, P., and Wilkinson, G. (1986). Towards a cartographic expert system. In *Proceedings AutoCarto*, London, UK, pages 578–587. [44](#)
- Mackaness, W. A. (2006). Automated cartography in a bush of ghosts. *Cartography and Geographic Information Science*, 33(4):245–256. [188](#), [189](#)
- Malone, L., Palmer, A. M., and Voigt, C. L. (2003). *Mapping Our World: GIS Lessons for Educators*. ESRI Press. [15](#)
- Mannes, J. (2004). Dynamische kartographische Visualisierung von Location Based Services mittels SVG. Master’s thesis, Department of Geography, Geographic Information Systems, University of Zurich, Switzerland. [xii](#), [38](#), [39](#)
- McMaster, R. B. (1989). The integration of simplification and smoothing algorithms in line generalization. *Cartographica*, 26(1):101–121. [43](#), [131](#), [132](#)
- McMaster, R. B. (1991). Conceptual frameworks for geographical knowledge. In Buttenfield, B. P. and McMaster, R. B., editors, *Map Generalization: Making Rules for Knowledge Representation*, Longman, London, UK, pages 21–39. [36](#)
- McMaster, R. B. (1995). Knowledge acquisition for cartographic generalization. In Müller, J. C., Lagrange, J. P., and Weibel, R., editors, *GIS and Generalization, GISDATA 1*, Taylor & Francis, pages 161–179. [44](#)
- McMaster, R. B. and Shea, K. S. (1992). *Generalization in Digital Cartography*. Association of American Geographers, Washington D.C. [44](#), [51](#), [70](#), [71](#), [91](#), [122](#)
- McMaster, R. B. and Usery, E. L. (2003). *A Research Agenda for Geographic Information Science*. Taylor & Francis. [12](#)
- Meier, S. (2000). Zur Qualität snakes-approximierter Höhenprofile mit Diskontinuitäten. *Photogrammetrie Fernerkundung Geoinformation*, 6:399–409. [131](#)

- Meng, L. (1997). Automatic generalization of geographic data. Technical report, VBB Viak, Stockholm, Sweden. [43](#)
- Meng, L. (2005). Egocentric design of map-based mobile services. *The Cartographic Journal*, 42(1):9–13. [17](#)
- Menke, K. (1982). *Zur rechnergestützten Generalisierung des Verkehrswege- und Gewässernetzes, insbesondere für den Maßstab 1:25'000*. PhD thesis, Universität Hannover, Germany. WAVH Nr. 119. [43](#)
- Müller, J. C., Weibel, R., Lagrange, J. P., and Salgé, F. (1995). Generalization: State of the art and issues. In Müller, J. C., Lagrange, J. P., and Weibel, R., editors, *GIS and Generalization. Methodology and Practice. GISDATA 1*, Taylor & Francis, pages 3–17. [34](#), [43](#), [51](#), [53](#)
- Morgenstern, D. and Schürer, D. (1999). A concept for model generalization of digital landscape models from finer to coarser resolution. In *Proceedings 19th International Cartographic Conference*, Ottawa, Canada, pages 1021–1028. [34](#), [45](#)
- Mountain, D., Goncalves, A., and Rodrigues, A. (2003). Geographic knowledge discovery and spatial agent interfaces for location-based services. In *Proceedings GIS Research UK 11th Conference*, London, UK, pages 206–209. [57](#), [180](#)
- Mountain, D. M. (2006). *Exploring mobile trajectories: An investigation of individual spatial behaviour and geographic filters for information retrieval*. PhD thesis, City University, London, UK. [19](#), [20](#), [57](#)
- Mustiere, S. (2005). Cartographic generalization of roads in a local and adaptive approach: A knowledge acquisition problem. *International Journal of Geographical Information Science*, 19(8-9):937–955. [45](#)
- Mustière, S. and Moulin, B. (2002). What is spatial context in cartographic generalisation? In *ISPRS & SIS, Geospatial Theory, Processing and Applications*, Ottawa, Canada, pages 274–278. [113](#), [138](#)
- Mustière, S., Saitta, L., and Zucker, J.-D. (2000). Abstraction in cartographic generalization. In *Proceedings of the 12th International Symposium on Foundations of Intelligent Systems*, Lecture Notes in Computer Science, Vol. 1932, Springer-Verlag, pages 638–644. [24](#)

- Mustière, S., Zucker, J.-D., and Saitta, L. (1999). Cartographic generalization as a combination of representing and abstracting knowledge. In *Proceedings of the 7th International Symposium on Advances in Geographic Information Systems, ACM-GIS*, Kansas City, USA, pages 162–164. [24](#)
- Nakos, B. and Miropoulos, V. (2003). Local length ratio as a measure of critical points detection for line simplification. In *Proceedings 6th Workshop on Progress in Automated Map Generalization*, Paris, France. CD-ROM. [122](#)
- Neun, M. and Burghardt, D. (2005). Web services for an open generalisation research platform. In *Proceedings 8th ICA Workshop on Generalisation and Multiple Representation*, A Coruña, Spain. CD-ROM. [40](#), [54](#)
- Neun, M., Burghardt, D., and Weibel, R. (2006). Spatial structures as generalisation support services. In *Joint ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data*, Hannover, Germany. CD-ROM. [55](#), [105](#), [183](#)
- Neun, M., Burghardt, D., and Weibel, R. (2008). Web service approach for providing enriched data structures to generalisation operators. *International Journal of Geographical Information Science*, 22(1-2):133–165. [55](#)
- Neun, M., Burghardt, D., and Weibel, R. (accepted). Automated processing for map generalization using web services. *GeoInformatica*. [55](#)
- Neun, M. and Steiniger, S. (2005). Modelling cartographic relations for categorical maps. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. [54](#), [102](#), [103](#)
- Neun, M., Weibel, R., and Burghardt, D. (2004). Data enrichment for adaptive generalisation. In *Proceedings 7th ICA Workshop on Generalisation and Multiple Representation*, Leicester, UK. CD-ROM. [71](#), [90](#), [102](#), [177](#)
- Nickerson, B. and Freeman, H. (1986). Development of a rule-based system for automatic map generalization. In *Proceedings of the Second International Symposium on Spatial Data Handling*, Seattle, Washington, USA, pages 537–556. [43](#), [44](#)
- Nivala, A.-M. and Sarjakoski, L. (2003a). An approach to intelligent maps: Context awareness. In *2nd Workshop on Human Computer Interaction in Mobile Guides*, Udine, Italy. [18](#)



- Nivala, A.-M. and Sarjakoski, L. T. (2003b). Need for context-aware topographic maps in mobile devices. In *Proceedings of Scandinavian Research Conference on Geographical Information Science*, Espoo, Finland, pages 15–29. 18
- Nivala, A.-M. and Sarjakoski, L. T. (2004). Preventing interruptions in mobile map reading process by personalisation. In *3rd Workshop on Human Computer Interaction in Mobile Guides*, Glasgow, Scotland. 20, 24
- OGC and ISO (2002). The OpenGIS Abstract Specification, Topic 12: OpenGIS Service Architecture, Version 4.3. Abstract Specification OGC 02-112. Also available as ISO/DIS 19119 - Geographic Information Services. Technical report, Open Geospatial Consortium. 60, 71
- Openshaw, S. and Alvandies, S. (1999). Applying geocomputation to the analysis of spatial distributions. In Goodchild, M. and Longley, P., editors, *Geographic Information Systems: Principles, Techniques, Management and Applications*, John Wiley and Sons Inc., New York, pages 267–282. 168
- Origel-Gutiérrez, G. (2004). Layout and design elements in the quality estimation of bi-dimensional internet based mapping. In *Proceedings of International Joint Workshop on Ubiquitous, Pervasive and Internet Mapping, UPIMap2004*, Tokyo, Japan. 15
- O’Sullivan, D. and Unwin, D. J. (2002). *Geographic Information Analysis*. John Wiley & Sons, Inc., New Jersey, USA. 167, 168, 174, 179
- Patalaviciute, V., Düpmeier, C., Freckmann, P., and Ruchter, M. (2005). Using SVG-based Maps for Mobile Guide Systems. A Case Study for the Design, Adaptation, and Usage of SVG-based Maps for Mobile Nature Guide Applications. In *SVG Open, 4th Annual Conference on Scalable Vector Graphics*, Enschede, Netherlands. xi, 20, 21, 24
- Perkal, J. (1966). An attempt at objective generalization. In *Michigan Inter-University Community of Mathematical Geographers*, Discussion Paper 10, pages 58–76. 43, 131
- Perlin, K. and Fox, D. (1993). Pad: an alternative approach to the computer interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH93*, ACM Press, pages 57–64. 31
- Peterson, M. P. (2006). *Maps and the Internet*. Elsevier Science Publication, 470 pages. 24



- Petzold, I. (2003). *Beschriftung von Bildschirmkarten in Echtzeit - Konzept und Struktur*. PhD thesis, Department of Cartography and Geodesy, University Bonn, Germany. [177](#)
- Petzold, I., Bobzien, M., and Burghardt, D. (2006a). Automatische Ableitung von Stadtplänen aus großmaßstäbigen Geodaten im Kontext des Projekts DRIVE. *Mitteilungen des Bundesamtes für Kartographie und Geodäsie*, 36:23–36. [56](#)
- Petzold, I., Burghardt, D., and Bobzien, M. (2005). Automated derivation of town plans from large scale data on an example of area to line simplification. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. [56](#)
- Petzold, I., Burghardt, D., and Bobzien, M. (2006b). Workflow management and generalisation services. In *Proceedings 9th ICA Workshop on Generalisation and Multiple Representation*, Vancouver, Washington. CD-ROM. [xii](#), [42](#), [55](#), [56](#)
- Plazanet, C. (1996). *Enrichissement des bases de données géographiques: analyse de la géométrie des objets linéaires pour la généralisation cartographique*. PhD thesis, Université de Marne-la-Vallée. [135](#)
- Plazanet, C., Bigolin, M., and Ruas, A. (1998). Experiments with learning techniques for spatial model enrichment and line generalization. *GeoInformatica*, 2(4):315–333. [122](#), [128](#)
- Powitz, B. M. (1988). Automationsgestützte Generalisierung von Verkehrswegedarstellungen. *Nachrichten aus dem Karten- und Vermessungswesen, Frankfurt a. M.*, 1(101). [44](#)
- Puppo, E. and Scopigno, R. (1997). Simplification, LOD and Multiresolution – Principals and Applications. In *EUROGRAPHICS 97, Tutorial Notes*. [140](#)
- Purves, R., Clough, P., Jones, C., Arampatzis, A., Bucher, B., Finch, D., Fu, G., Joho, H., Khirini, A., Vaid, S., and Yang, B. (2007). The design and implementation of spirit: a spatially-aware search engine for information retrieval on the internet. *International Journal of Geographical Information Science*, 21(7):717 – 745. [13](#)
- Purves, R., Syed, A. K., Yang, B., and Weibel, R. (2005). A cartographic visualisation interface for spatial information retrieval. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. [30](#)
- Ramer, U. (1972). An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1:244–256. [43](#)

- Raper, J., Dykes, J., Wood, J., Mountain, D., Krause, A., and Rhind, D. (2002). A framework for evaluating geographical information. *Journal of Information Science*, 28(2):39–50. 157, 158, 160
- Rappo, A. (2003). Fischaugenprojektionen mit distanzabhängiger Informationsdichte für die kartographische Visualisierung auf kleinen Bildschirmen. Master's thesis, Department of Geography, University of Zurich, Switzerland. xi, 31, 32
- Rappo, A., Cecconi, A., and Burghardt, D. (2004). Fischaugenprojektionen für generalisierte Kartendarstellungen auf kleinen Bildschirmen. *Kartographische Nachrichten*, 2:73–79. 31
- Ratajski, L. (1967). Phénomène des points de généralisation. *International Yearbook of Cartography*, 7:143–152. 43
- Regnauld, N. (1996). Recognition of building clusters for generalization. In *Proceeding of the 7th International Symposium on Spatial Data Handling*, Delft, Netherlands, pages 4B.1–4B.14. 138, 175
- Regnauld, N. (2001). Constraint based mechanism to achieve automatic generalization using agent modelling. In *Proceedings GIS Research UK 9th Annual Conference*, Glamorgan, Wales, pages 329–332. 37, 46
- Regnauld, N. (2006). Improving efficiency for developing automatic generalisation solutions. In *Joint ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data*, Hannover, Germany. 40
- Reichenbacher, T. (2001). The world in your pocket - towards a mobile cartography. In *Proceedings 20th International Cartographic Conference*, Beijing, China, pages 2514–2522. 16
- Reichenbacher, T. (2004). *Mobile Cartography - Adaptive Visualisation of Geographic Information on Mobile Devices*. PhD thesis, Lehrstuhl für Kartographie der Universität München. Verlag Dr. Hut, Munich, Germany. 20, 174
- Reumann, K. and Witkam, A. P. M. (1974). Optimizing curve segmentation in computer graphics. In *International Computing Symposium*, Amsterdam, Netherlands, pages 467–472. 43
- Richardson, D. E. (1989). Rule based generalization for base map production. In *Proceedings of Challenge for the 1990s: Geographic Information Systems Conference*, Ottawa, Canada, pages 718–739. 44

- Robinson, G. and Zaltash, A. (1989). Application of expert systems to topographic map generalisation. In *Proceedings of the Association for Geographic Information Conference*, pages A.3.1–A.3.6. [44](#)
- Ruas, A. (1998). Method for building displacement in automated generalisation. *International Journal of Geographical Information Science*, 12(8):789–803. [172](#)
- Ruas, A. (1999). *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*. PhD thesis, Université de Marne-la-Vallée. [37](#), [46](#), [83](#)
- Ruas, A. (2000). The roles of meso objects for generalization. In *Proceedings, 9th Symposium on Spatial Data Handling*, Beijing, China, pages 3b50–3b63. [70](#), [71](#), [104](#), [135](#)
- Ruas, A. (2001). Automatic generalisation project: Learning process from interactive generalisation. Technical report, European Organization for Experimental Photogrammetric Research, Official Publication No. 39. [52](#)
- Ruas, A., editor (2002). *Généralisation et représentation multiple*. Hermès Science Publications. [36](#), [51](#)
- Ruas, A. and Holzapfel, F. (2003). Automatic characterisation of building alignments by means of expert knowledge. In *Proceedings of 21st International Cartographic Conference*, Durban, South Africa, pages 1604–1615. [87](#), [104](#), [107](#)
- Ruas, A., Mackaness, W. A., and Sarjakoski, T., editors (2007). *Generalisation of Geographic Information: Cartographic Modelling and Applications*. Series of International Cartographic Association, Elsevier Science. [51](#)
- Ruas, A. and Plazanet, C. (1996). Strategies for automated generalization. In *Proceedings 7th International Symposium on Spatial Data Handling (=Advances in GIS Research II)*, Taylor & Francis, Delft, Netherlands, pages 6.1–6.17. [90](#)
- Sako, Y. and Fujimura, K. (2000). Shape similarity by homotopic deformation. *The Visual Computer*, 16(1):47–61. [30](#)
- Samet, H. (1990). *Applications of spatial data structures*. Addison-Wesley: New York. [167](#)
- Sarjakoski, T., Sester, M., Sarjakoski, L., Harrie, L., Hampe, M., Lehto, L., and Koivula, T. (2005). Web Generalisation Service in GiMoDig - towards a standardised service for real-time generalisation. In *Proceedings of the 8th AGILE Conference on GI-Science*, Estoril, Portugal. [22](#), [38](#)

- Saux, E. (2003). B-spline functions and wavelets for cartographic line generalization. *Cartography and Geographic Information Science*, 30(1):33–50. [131](#)
- Schapire, R. E. (1999). A brief introduction to boosting. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, Stockholm, Sweden, pages 1401–1406. [96](#)
- Scharlach, H. and Müller, J. (2001). Multimediale thematische Kartographie auf Handheld PCs: Potentiale und Grenzen. In *Proceedings of 2nd Symposium on Location Based Services and TeleCartography*, Wien, Austria, pages 87–98. [17](#)
- Schittenhelm, R. (1980). Ein Programm zur EDV-unterstützten Entflechtung der Überlappungsstellen verbreiteter linienhafter Kartenelemente. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(80):75–82. [43](#)
- Schlieder, C., Vögele, T., and Werner, A. (2001). Location modeling for intentional behavior in spatial partonomies. In *Proceedings of the Workshop on Location Modeling for Ubiquitous Computing*, Atlanta, Georgia, USA, pages 63–70. [161](#), [162](#)
- Schroeder, W. J., Zarge, J. A., and Lorensen, W. E. (1992). Decimation of triangle meshes. *Computer Graphics*, 26:65–70. [138](#)
- Schuurman, N. (1999). Critical GIS: Theorizing an emerging science. *Cartographica Vol. 36*, 36(4):7–10. [110](#)
- Schwarzbach, F. (1995). *Untersuchungen zur rechnergestützten Linienglättung*. PhD thesis, Department of Cartography, University of Dresden. Kartographische Bausteine 10. [131](#)
- Schweinfurth, G. (1984). *Höhenliniengeneralisierung mit Methoden der digitalen Bildverarbeitung*. PhD thesis, Fakultät für Bauingenieur- und Vermessungswesen, Universität Karlsruhe, Deutsche Geodätische Kommission, Reihe C, No. 291. [131](#), [132](#)
- Schylberg, L. (1993). *Computational Methods for Generalization of Cartographic Data in a Raster Environment*. PhD thesis, Royal Institute of Technology, Stockholm. [44](#)
- Sederberg, T. W. and Scott, R. P. (1986). Free-form deformation of solid geometric models. In *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, ACM-Press, New York, pages 151–160. [30](#)
- Sester, M. (2000). Generalization based on least squares adjustment. In *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B4, Amsterdam, Netherlands, pages 931–938. [46](#), [71](#), [123](#), [175](#)

- Sester, M., Anders, K.-H., and Walter, V. (1998). Linking objects of different spatial data sets by integration and aggregation. *GeoInformatica*, 2(4):335–358. 113
- Sester, M. and Brenner, C. (2000). Typification based on kohonen feature nets. In: *Proceedings of the 1st International Conference on GIScience, Savannah (GA)*., pages 21–22. 138
- Sester, M. and Brenner, C. (2004a). Continuous generalization for fast and smooth visualization on small displays. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 35, Istanbul, Turkey, pages 1293–1298. 27
- Sester, M. and Brenner, C. (2004b). Continuous generalization for visualization on small mobile devices. In Fisher, P., editor, *Developments in Spatial Data Handling, 11th International Symposium on Spatial Data Handling*, Springer Verlag, pages 469–480. 27
- Sester, M., Sarjakoski, L. T., Harrie, L., Hampe, M., Koivula, T., Sarjakoski, T., Lehto, L., Elias, B., Nivala, A.-M., and Stigmar, H. (2004). Real-time generalization and multiple representation in the gimodig mobile service. Technical report, GiMoDig-project, IST-2000-30090, Deliverables D7.1.1\*, D7.2.1\* and D7.3.1, Public EC report, 151 pages. 75
- SGK (2002). *Topografische Karten, Kartengrafik und Generalisierung*. Kartographische Publikationsreihe Nr. 16, Schweizerische Gesellschaft für Kartographie. CD-ROM. 173
- Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages*, Washington, USA, IEEE Computer Society Press, pages 336–343. 12
- Shreve, R. L. (1966). Statistical law of stream number. *Journal of Geology*, 74:17–37. 43
- Slocum, T. (1999). *Thematic Cartography and Visualization*. Prentice Hall, Upper Saddle River, New Jersey. 172
- Spaccapietra, S., Parent, C., and Vangenot, C. (2000). GIS databases: From multiscale to multirepresentation. In Choueiry, B. and Walsh, T., editors, *Proceedings of the 4th International Symposium on Abstraction, Reformulation, and Approximation*, Vol. 1864, Lecture Notes in Artificial Intelligence, Springer, pages 57–70. 45
- Spiess, E. (1990). Siedlungsgeneralisierung. *Kartographische Publikationsreihe*, 10. Schweizerische Gesellschaft für Kartographie, Zürich. 150

- SSC (2005). *Topographic Maps - Map Graphics and Generalization*. Cartographic Publication Series, Vol. 17 (Swiss Society of Cartography: Bern). CD-ROM. [xiii](#), [26](#), [102](#), [108](#), [167](#)
- Staufenbiel, W. (1974). Zur Automation der Siedlungsgeneralisierung unter besonderer Berücksichtigung der Formvereinfachung. *Nachrichten aus dem Karten- und Vermessungswesen*, 1(65):101–110. [43](#)
- Steiniger, S., Lange, T., Weibel, R., and Burghardt, D. (2007). An approach for the classification of urban building structures based on discriminant analysis techniques. *Transaction in GIS*. in press. [54](#), [184](#)
- Steiniger, S. and Meier, S. (2004). Snakes: A technique for line smoothing and displacement in map generalisation. In *Proceedings 7th ICA Workshop on Generalisation and Multiple Representation*, Leicester, UK. CD-ROM. [129](#)
- Steiniger, S. and Weibel, R. (2005a). A conceptual framework for automated generalization and its application to geologic and soil maps. In *Proceedings of 22nd International Cartographic Conference*, A Coruña, Spain. CD-ROM. [91](#)
- Steiniger, S. and Weibel, R. (2005b). Relations and structures in categorical maps. In *Proceedings 8th ICA Workshop on Generalisation and Multiple Representation*, A Coruña, Spain. CD-ROM. [93](#)
- Steiniger, S. and Weibel, R. (2007). Relations between map objects in cartographic generalization. *Cartography and Geographic Information Science*. in press. [103](#)
- Steiniger, S., Weibel, R., and Burghardt, D. (2006a). Recognition of large island structures for map generalization. In Raubal, M., Miller, H. J., Frank, A. U., and Goodchild, M., editors, *4th International Conference, Geographic Information Science*, IfGIprints 28, pages 351–354. [54](#)
- Steiniger, S., Weibel, R., and Burghardt, D. (2006b). Recognition of perceptual island structures for map generalization. In *Proceedings of 14th International Symposium on Advances in Geographic Information Systems*, Arlington (VA), USA, pages 67–74. [54](#)
- Stoter, J., Anders, K.-H., Duchêne, C., Baella, B., Burghardt, D., Dortland, A., Pla, M., Regnauld, N., Rosenstand, P., Storm, M., and Touya, G. (2006). EuroSDR project – Research project on state-of-the-art of generalisation. Technical report, European Spatial Data Research. [53](#)



- Stoter, J. E. (2005). Generalisation: The gap between research and practice. In *Proceedings 8th ICA Workshop on Generalisation and Multiple Representation*, A Coruña, Spain. CD-ROM. 52
- Strahler, A. N. (1952). Hypsometric (area-altitude) analysis of erosional topology. *Bulletin of the Geological Society of America*, 63:1117–1142. 43, 162
- Summerfield, M. A. (1991). *Global Geomorphology*. Longman: London, UK. 162, 168, 179
- Tezuka, T. and Tanaka, K. (2005). Trajectory-based presentation of heterogeneous spatio-temporal content. In *5th International Workshop on Web and Wireless Geographical Information Systems*, Lecture Notes in Computer Science, Springer, pages 235–245. 19
- Thomas, J. J. and Cook, K. A., editors (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Center, IEEE CS Press. 11
- Timpf, S. (1997). Cartographic objects in a multi-scale data structure. In Craglia, M. and Couclelis, H., editors, *Geographic Information Research: Bridging the Atlantic*, Taylor & Francis, London, UK, pages 224–234. 167
- Timpf, S. (1998). *Hierarchical structures in map series*. PhD thesis, Technical University Vienna, Austria. 105
- Tobler, W. (2004). Thirty five years of computer cartograms. *Annals of the Association of American Geographers*, 94:58–73. 30
- Tobler, W. R. (1964). An experiment in the computer generalization of maps. Technical report, Office of Naval Research, ASAD Contract 459953, Washington, DC.  
Taken from: Battenfield, B. and McMaster, R. (1991). *Map Generalization - Making Rules for Knowledge Representation*. Longman Group, UK. 43
- Tobler, W. R. (1966). Numerical map generalization. In *Michigan Inter-University Community of Mathematical Geographers*, Discussion Paper 8.  
Taken from: Battenfield, B. and McMaster, R. (1991). *Map Generalization - Making Rules for Knowledge Representation*. Longman Group, UK. 43
- Tobler, W. R. (1973). A continuous transformation useful for districting. *Annals of the New York Academy of Sciences*, 219(9):215–220. 30



- Tobler, W. R. (1986). Pseudo-cartograms. *The American Cartographer*, 13(1):43–50. 30
- Töpfer, F. (1974). *Kartographische Generalisierung, Ergänzungsheft Nr. 276 zu Petermanns Geographische Mitteilungen*. VEB Hermann Haack, Geographisch-Kartographische Anstalt Gotha/Leipzig. xiv, 51, 123
- Töpfer, F. and Pillewitzer, W. (1966). The principle of selection, a means of cartographic generalisation. *The Cartographic Journal*, 3(1):10–16. 43, 144
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading MA: Addison-Wesley. 11
- Turk, G. (1992). Re-tiling polygonal surfaces. *Computer Graphics*, 26:55–64. 138
- Turesson, H. and Harrie, L. (2003). Integration of navigational and cartographic data. In *Proceedings of 21st International Cartographic Conference*, Durban, South Africa, pages 2572–2578. 38
- van Kreveld, M. (2001). Smooth generalization for continuous zooming. In *Proceedings 20th International Cartographic Conference*, Beijing, China, pages 2178–2185. 27, 28
- van Oosterom, P. (1989). A reactive data structure for geographic information systems. In *Proceedings of AutoCarto 9*, Baltimore, Maryland, USA, pages 665–674. 177
- van Oosterom, P. (1995). The GAP-tree, an approach to ‘on-the-fly’ map generalization of an area partitioning. In Müller, J. C., Lagrange, J. P., and Weibel, R., editors, *GIS and Generalization, GISDATA 1*, Taylor & Francis, pages 120–132. 29, 45, 177
- Vassiliadis, P., Simitsis, A., and Skiadopoulos, S. (2002). Conceptual modeling for ETL processes. In *Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP*, ACM Press, New York, NY, USA, pages 14–21. 158
- W3C (2004). Web services architecture. Technical report, <http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/>. 61
- Waldeck, C. and Balfanz, D. (2004). Mobile Liquid 2D Scatter Space (ML2DSS). In Banissi, E., editor, *Proceedings 8th International Conference on Information Visualization*, Los Alamitos, California, USA, pages 494–498. 31

- Ware, J., Jones, C., and Thomas, N. (2003). Automated map generalization with multiple operators: A simulated annealing approach. *International Journal of Geographical Information Science*, 17(8):743–769. [46](#), [71](#)
- Weakliam, J., Bertolotto, M., and Wilson, D. (2005). Implicit interaction profiling for recommending spatial content. In Shahabi, C. and Boucelma, O., editors, *13th ACM International Workshop on Geographic Information Systems, ACM-GIS*, Bremen, Germany, pages 285–294. [18](#)
- Weibel, R. (1991). Amplified intelligence and rule-based systems. In Buttenfield, B. P. and McMaster, R. B., editors, *Map Generalization: Making Rules for Knowledge Representation*, Longman, London, pages 172–186. [44](#)
- Weibel, R. (1993). Knowledge acquisition for map generalization: Methods and prospects. In *Specialist Meeting on Formalizing Cartographic Knowledge*, Buffalo, NY, USA, pages 223–232. [44](#)
- Weibel, R. (1995). Three essential building blocks for automated generalization. In Müller, J. C., Lagrange, J. P., and Weibel, R., editors, *GIS and Generalization, GISDATA 1*, Taylor & Francis, pages 56–69. [44](#)
- Weibel, R. (1996). A typology of constraints of line simplification. In *Proceedings 7th International Symposium on Spatial Data Handling (=Advances in GIS Research II)*, Delft, Netherlands, Taylor & Francis, pages 9A.1–9A.14. [83](#)
- Weibel, R. (1997). Generalization of spatial data principles and selected algorithms. In van Kreveld, M., Nievergelt, J., Roos, T., and Widmayer, P., editors, *Algorithmic Foundations of Geographic Information Systems*, Lecture Notes in Computer Science, Vol. 1340, Springer Verlag, pages 99–152. [43](#), [51](#), [160](#)
- Weibel, R. and Burghardt, D. (2007). On-the-fly generalization. In Kemp, K., editor, *Encyclopedia of Geographic Information Science*, Sage Publications. [22](#), [27](#)
- Weibel, R. and Dutton, G. (1998). Constraint-based automated map generalization. In *Proceeding 8th Int. Symp. on Spatial Data Handling (SDH)*, pages 214–224. [xii](#), [34](#), [35](#)
- Weibel, R. and Dutton, G. (1999). Generalising spatial data and dealing with multiple representations. In Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W., editors, *Geographical Information Systems*, Vol. 1, John Wiley & Sons, Inc., pages 125–155. [45](#)

- Weibel, R. and Jones, C. B. (1998). Computational perspective on map generalization. *GeoInformatica*, 2(4):307–314. 172
- Weibel, R., Keller, S., and Reichenbacher, T. (1995). Overcoming the knowledge acquisition bottleneck in map generalization: The role of interactive systems and computational intelligence. In Frank, A. U. and W. Kuhn, W., editors, *Spatial Information Theory, International Conference COSIT95*, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, pages 139–156. 44
- Weissenberg, N., Gartmann, R., and Voisard, A. (2006). An ontology based approach to personalized situation aware mobile service supply. *GeoInformatica*, 10(1):55–90. 18
- Werschlein, T. and Weibel, R. (1994). Use of neural networks in line generalization. In *Proceedings of the 5th European Conference and Exhibition on Geographic Information Systems, EGIS'94*, Paris, France, pages 76–85. 44
- Wertheimer, M. (1938). Laws of organization in perceptual forms. In Ellis, W., editor, *A Source Book of Gestalt Psychology*, London, Routledge & Kegan Paul, pages 71–88. 43
- WFS (2005). Opengis web feature service interfaces implementation specification. Technical report, <http://www.opengeospatial.org/specs/?page=specs>. 60
- White, R. W. (2005). *Implicit Feedback for Interactive Information Retrieval*. PhD thesis, Department of Computing Science, University of Glasgow, Scotland. 18
- Wilson, I. D., J. Mark Ware, J. M., and Ware, J. A. (2003). A genetic algorithm approach to cartographic map generalisation. *Computers in Industry*, 52(3):291–304. 46
- Winter, S. and Tomko, M. (2004). Shifting the focus in mobile maps. In *Proceedings Workshop on Ubiquitous, Pervasive and Internet Mapping, UBIMap*, Tokyo, Japan, pages 153–165. 17
- WMS (2005). OpenGIS Web Map Service interfaces implementation specification. Technical report, <http://www.opengeospatial.org/specs/?page=specs>. 60
- Yang, B., Purves, R., and Weibel, R. (2007). Efficient transmission of vector data over the internet. *International Journal of Geographical Informatin Science*, 21(1-2):215–237. xi, 29

- Zhang, J., Zhu, M., Papadias, D., Tao, Y., and Lee, D. (2003). Location-based spatial queries. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, ACM Press, New York, USA, pages 443 – 454. [20](#)
- Zhang, Q. and Harrie, L. (2006). Real-time map labelling for mobile applications. *Computers, Environment and Urban Systems*, 30(6):773–783. [22](#)
- Zoraster, S. (1986). Integer programming applied to the map label placement problem. *Cartographica*, 23(3):16–27. [45](#)
- Zoraster, S. (1997). Practical results using simulated annealing for point feature label placement. *Cartography and Geographic Information Systems*, 24(4):228–238. [45](#)
- Zuendorf, G., Kerrouche, N., Herholz, K., and Baron, J. C. (2003). Efficient principal component analysis for multivariate 3D voxel-based mapping of brain functional imaging data sets as applied to FDG-PET and normal aging. *Human Brain Mapping*, 18:13–21. [98](#)